3
Noise

3.1 INTRODUCTION

The noise components of a signal can have different origins. Sometimes noise is human-made (e.g., artifacts from switching instruments or 60-Hz hum originating from power lines). Other noise sources are random in nature, such as thermal noise originating from resistors in the measurement chain. Random noise is intrinsically unpredictable, but it can be described by statistics. From a measurement point of view, we can have noise that is introduced as a result of the measurement procedure itself, either producing ***systematic bias*** (e.g., measuring the appetite after dinner) or random ***measurement noise*** (e.g., thermal noise added by recording equipment). If we consider a measurement M as a function of the measured process x and some additive noise N, the *ith* measurement can be defined as

$$M_i = x_i + N_i \tag{3.1}$$

An example with $x_i = 0.8x_{i-1} + 3.5$ plus the noise contribution drawn from a random process is shown in Figure 3.1A. This trace was produced by pr3_1.m.

Alternately, noise may be intrinsic to the process under investigation. This ***dynamical noise*** is not an independent additive term associated with the measurement but instead interacts with the process itself. For example, temperature fluctuations during the measurement of cellular membrane potential not only add unwanted variations to the voltage reading; they physically influence the actual processes that determine the potential. If we consider appropriately small time steps, we can imagine the noise at one time step contributing to a change in the state at the next time step. Thus, one way to represent dynamical noise D affecting process x is

$$x_i = [0.8x_{i-1} + 3.5] + D_{i-1} \tag{3.2}$$

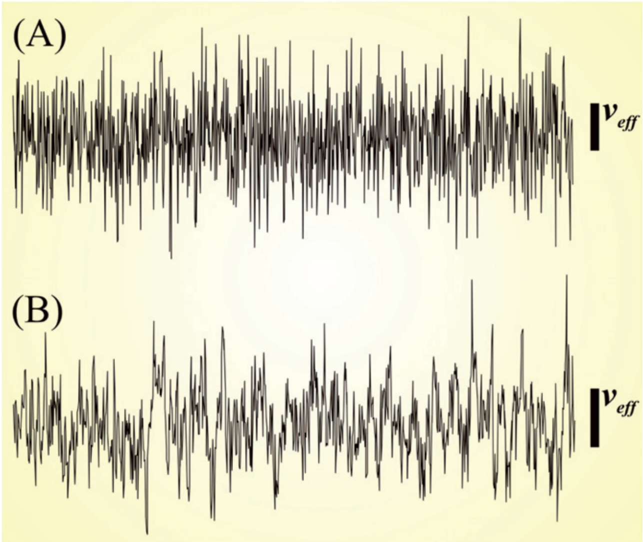

Figure 3.1 Time series including measurement noise (A) and a combination of dynamical and measurement noise (B). These examples were generated with MATLAB scripts pr3_1 and pr3_2. The bars on the right side represent the v_{eff} level for each signal (Equation (3.14)).

The process in Equation (3.2) can be combined with a measurement function such as Equation (3.1). Comparing the time series of such a process (Fig. 3.1B, generated by pr3_2.m) with the one generated by Equation (3.1), you can see that the dynamical noise (due to the correlation between sequential values) creates slower trends when compared to the time series with only additive noise. It must be noted here that in many cases, a dynamic noise term is used to represent a random process simply because often we do not know all of the details necessary to accurately represent the entire range of complex interactions in a physiological system. In this sense, the random process compensates for our lack of detailed knowledge by giving us a statistical proxy for what we do not know about the system. As we will see in the discussion of nonlinear dynamics (Chapter 17) ***deterministic*** processes (processes in which the state is ***determined*** by the past) can produce signals with a random aspect — that is, in some cases the difference between the behavior of a random number generator and a deterministic process can become fuzzy. These processes are similar to the bouncing balls in a lotto drawing; while the outcome is ultimately the result of completely deterministic physical laws, the exact result is entirely unpredictable.

Note: The process in Equation (3.1) is deterministic; only its measurement is corrupted by noise. However, although the process in Equation (3.2) includes a deterministic component, it is a so-called stochastic process because a noise component is part of the process itself.

3.2 NOISE STATISTICS

One common way to characterize a random process is by its ***probability density function (PDF)***, describing the probability $p(x)$ that particular values of $x(t)$ occur. For instance, if we create a function to describe the probability of each outcome of a fair roll of a single die, we would have the possible observations 1, 2, 3, 4, 5, and 6. In this case, each of the six possible observations occurs with a probability $p(1)$, $p(2)$, . . . , $p(6)$, each equal to one sixth. This would result in a PDF that is 1/6 for each of the values 1 through 6 and 0 for all other values. The PDF for the fair die is shown in Figure 3.2A. This example can be extended to continuous variables, and such an example of a variable that ranges between 0 and 6 is shown in Figure 3.2B. In this example, all values within the range are equally likely to occur. Often this is not the case; the most well-known PDF is the normal distribution shown in Figure 3.2C, reflecting a process where most values are close to the mean and extreme values (either positive or negative) are less likely to occur.

Note: The function describing the probability function of a discrete random variable is often called the ***probability mass function (PMF)***. In this text, we use the term *probability density function* both in the case of discrete and continuous random variables.

In general, a PDF characterizes the probabilities of all possible outcomes of random event, so the sum of the probabilities must equal 1, and the component probability values are therefore fractions less than 1. In the case of the single die, the total is

$$p(1)+p(2)+p(3)+p(4)+p(5)+p(6)=\sum_{i=1}^{6}p(i)=1, \quad \text{with } p(i)=1\div 6$$

In the case of continuous random variables, we replace the summation by an integral over the domain of x, which translates intuitively into the requirement that the area under the PDF must equal 1. In the case of a continuous uniform distribution as in Figure 3.2B, we integrate over the domain 0 to 6 — that is, $\int_0^6 p(x)\,dx=1$. More generally, as in the example in Figure 3.2C, we consider a domain from $-\infty$ to ∞:

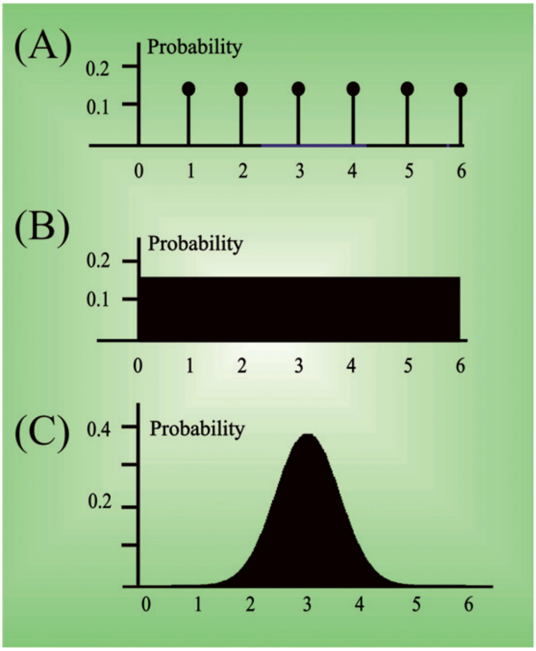

Figure 3.2 Probability density functions (PDF) of random processes. (A) The PDF of a die where each of the outcomes 1 to 6 is equally likely. (B) A similar uniform distribution for a continuous process. An example of such a process is quantization noise caused by analog-to-digital conversion (see Section 3.4.4). (C) The normal distribution, where probabilities are not uniform across the domain. Values close to the mean are more likely to occur as compared to more extreme values. In this example, the mean of the normal distribution is 3, while the standard deviation and variance are both equal to 1.

$$\int_{-\infty}^{\infty} p(x)\,dx = 1 \tag{3.3}$$

Two useful variations on the PDF can be derived directly from it: the ***cumulative*** $F(x)$ and ***survival*** $F(x)$ functions are defined as

$$\mathcal{F}(x) = \int_{-\infty}^{x} p(y)\,dy \tag{3.4}$$

$$\mathcal{F}(x) = 1 - F(x) = \int_{x}^{\infty} p(y)\,dy \tag{3.5}$$

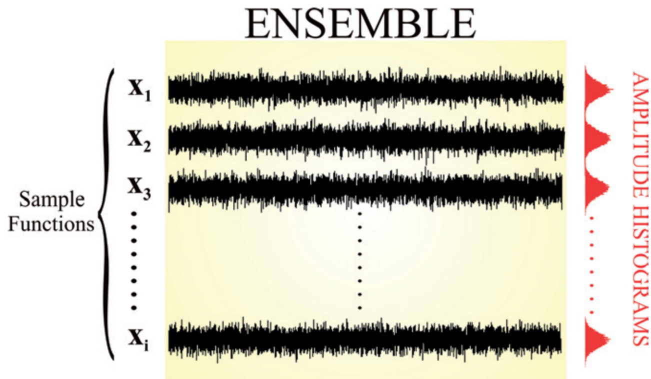

Figure 3.3 Observations of the random process characterized by the PDF shown in Figure 3.2C. Sample functions are individual "samples" from the larger ensemble. For each trace, the amplitude distribution histogram is shown on the side in red. To present amplitude in both the sample functions and histograms along the same axis, the orientation of the amplitude distribution histogram is rotated 90 degrees from that used in Figure 3.2C (i.e., the vertical axis of this distribution corresponds to the range of amplitude values and the horizontal axis to the number of times this amplitude was present in the associated sample function).

As can be inferred from the integration limits in Equations (3.4) and (3.5), the cumulative function $(-\infty, x)$ represents the probability that the random variable is $\leq x$, and the survival function (x, ∞) represents $p(y) > x$.

If one observes a random process over time, one can obtain sample functions, series of measured values representing one instance of the random process (Fig. 3.3). A collection of these sample functions forms an ***ensemble***. The random process is called ***stationary*** if the distribution from which $x(t)$ originated does not change over time. In Figure 3.3, the amplitude distribution is shown for each sample function. The similarity of these distributions makes the assumption of underlying stationarity a reasonable one. The process is ***ergodic*** if any of the particular sample functions is representative of the whole ensemble, thus allowing statistics to be obtained from averages over time. When applying signal processing techniques, the stationarity and ergodicity of signals are frequently (and implicitly) assumed, and many techniques can be useful even when these assumptions are not strictly met. Other, less stringent, definitions for both terms also exist (Appendix 3.1).

Two common parameters that are estimated from random processes are mean and variance. If a process is stationary and ergodic, one can characterize the distribution using any of the sample functions (Fig. 3.1) — that is, the ***estimate*** of the mean of x over an interval T is

$$\bar{x} = \frac{1}{T}\int_0^T x(t)\,dt \tag{3.6}$$

or for a discrete-valued signal over N points:

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{3.7}$$

Similarly, one can estimate the variance from the time series:

$$\overline{Var(x)} = \frac{1}{N}\sum_{i=1}^{N} (x_i - \bar{x})^2 \tag{3.8}$$

To obtain a nonbiased estimate of the variance with small samples, $N - 1$ instead of N is used in the denominator of the scaling term. In the previous approach to estimating statistics from a sample of an ergodic process, a value close to the *true mean* $\langle x \rangle$ is obtained as the interval T extends toward infinity:

$$\langle x \rangle = \lim_{T\to\infty} \frac{1}{T}\int_0^T x(t)\,dt$$

A different approach to obtaining the true mean and standard deviation is via the probability density function (PDF) of the observed variable x, using the *Expectation* $E\{x\}$:

$$E\{x\} = \int_{-\infty}^{\infty} x\,p(x)\,dx = \langle x \rangle \tag{3.9}$$

In general, one can use the expectation to obtain the *nth* ***moment*** of the distribution:

$$E\{x^n\} = \int_{-\infty}^{\infty} x^n p(x)\,dx \tag{3.10}$$

or the *nth central moment*:

$$E\{(x - \langle x \rangle)^n\} = \int_{-\infty}^{\infty} (x - \langle x \rangle)^n\, p(x)\,dx \tag{3.11}$$

The first moment is the ***mean*** **(*μ*)**, the second central moment is the *variance* (σ^2), and the square root of the variance is the ***standard deviation*** **(*σ*)**. The square root of the variance of the estimate of the mean is the ***standard error of the mean*** (***SEM;*** see Chapter 4). The first central moment of a joint distribution of two variables, x and y, is the ***covariance*** — that is, $E\{(x - \langle x \rangle)(y - \langle y \rangle)\}$.

Note: The Laplace and Fourier transforms of the PDFs are sometimes used to generate the moments of the distribution (Appendix 3.4).

3.3 SIGNAL-TO-NOISE RATIO

Generally, any (biomedical) measurement will necessarily be corrupted by some noise. Even if the process itself were noise free, the measurement chain adds noise components because all analog instruments (amplifiers, analog filters) add, at the very least, a small amount of thermal noise (e.g., Equation (3.1)). If the noise component is sufficiently small compared to the signal component, one can still gather reasonable measurements of the signal. To quantify this ratio between signal and noise components, one can (in some cases) determine the amplitude or the power of each component and from those calculate a ***signal-to-noise ratio***. In discrete time series, the *power* can be measured as the mean squared amplitude $\left(\boldsymbol{ms}, \frac{1}{N}\sum_{i=1}^{N} x_i^2\right)$ and the *amplitude* as the root of the mean squared amplitude $\left(\boldsymbol{rms}, \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2}\right)$. Analytical equivalents for continuous time series are $ms = \frac{1}{T}\int_0^T x(t)^2\, dt$, and the *rms* is $\sqrt{\frac{1}{T}\int_0^T x(t)^2\, dt}$. To establish the signal-to-noise ratio (*SNR*), one can use $\frac{ms(signal)}{ms(noise)}$ directly; however, it is more common to represent this ratio on a logarithmic decibel (dB) scale:

$$SNR = 10\log_{10}\frac{ms(signal)}{ms(noise)}\text{dB} \tag{3.12}$$

Alternatively, one may start from the *rms* values by substituting $ms = rms^2$ in Equation (3.12):

$$SNR = 10\log_{10}\left[\frac{rms(signal)}{rms(noise)}\right]^2 = 20\log_{10}\frac{rms(signal)}{rms(noise)}\text{dB} \tag{3.13}$$

Note that the dB scale does not have a physical dimension; it is simply the logarithm of a ratio. The signal-to-noise ratio (without the log transform) is sometimes used as a *figure of merit* (*FOM*) by equipment manufacturers. If this ratio is close to 1, or even less than 1, signal processing can help to increase SNR in special cases.

In technical literature for analog devices, the noise level of $v(t)$ in an interval T is frequently indicated with v_{eff}, which equals the standard deviation of the signal:

$$v_{eff} = \sqrt{\frac{1}{T}\int_0^T (v-\overline{v})^2\, dt} \tag{3.14}$$

In the case of a sampled signal, the equivalent would be $\sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i-\overline{x})^2}$, similar to the definition of rms presented earlier.

> *Note:* To obtain a better looking figure for the noise specification, most manufacturers present v_{eff} after it has been corrected for any amplification. For instance, if a 1000× amplifier has 1 mV effective noise, a v_{eff} of 1 *μV at the input* is reported.

For noise with a zero mean, v_{eff} is the square root of $E\{x^2\}$; in this case, the difference between v_{eff} and *rms* disappears! It should further be noted that when observing a noise signal on a scope or chart writer, the ***amplitude of the noise band one observes is typically 4 to 5 times the*** v_{eff} (Fig. 3.1). The effects of combined noise sources add up geometrically in the total result: the total v_{eff} of two ***independent*** noise sources 1 and 2 in series, such as the noise generated in two connected instruments in a measurement chain, can be found by

$$v_{eff} = \sqrt{\left(v_{eff,1}^2 + v_{eff,2}^2\right)} \tag{3.15}$$

In MATLAB you can verify this by creating two random time series (s1 and s2) and the total result (st) by typing the following in the command window:

```
s1 = randn(1000, 1);
s2 = randn(1000, 1);
st = s1 + s2;
```

You will find that the v_{eff}^2 (variance) of st (vt) will be close to the sum of variances of s1 (v1) and s2 (v2); for example type

```
v1 = (std(s1))^2
v2 = (std(s2))^2
vt = (std(st))^2
```

Due to the random aspect of the time series, the outcome of this little numerical experiment will be a bit different each time, but in each case you will find that vt ≈ v1 + v2.

3.4 NOISE SOURCES

In the measurement chain there are several sources of noise, and some of these sources can be extremely annoying for the experimenter. The following summarizes four major sources of noise in the measurement chain discussed in Chapter 2.

1. Thermal or Johnson noise originating from resistors in the circuitry. The value can be estimated by

$$v_{eff}^2 = 4kTR\Delta f \tag{3.16}$$

k = 1.38 10^{-23}, T absolute temperature (°K), R resistor value, and Δf bandwidth.

Problem

Calculate v_{eff} of the noise generated by a Giga seal (10^9 Ω) made between a patch clamp electrode and a neuron. Assume a temperature of 27°C and a recording bandwidth of 10 kHz.

Answer

Using Equation (3.16) taking into account the conversion from °C into °K (by adding 273) we get

$$v_{eff}^2 = 4 \times 1.38 \times 10^{-23} \times (27 + 273) \times 10^9 \times 10^4 = 1.6560 \times 10^{-7}\ \mathrm{V}^2$$

Taking the square root of the outcome we find $v_{eff} \approx$ 0.4 mV.

Usually thermal noise is associated with a particular application, and it is rarely under direct control in a given setup. There are cases where designers have included cooling of the preamplifier (using a Peltier element as cooling device) to reduce thermal noise from the input resistors. The usefulness of this approach is limited because the temperature factor in Equation (3.14) is in °K, where a decrease of 10 degrees only reduces v_{eff} by a few percentage points.

2. Finding sources of ***(a) electromagnetic*** or ***(b) electrostatic*** noise (usually hum from power lines) can be a frustrating exercise. Generally, noise caused by a fluctuating magnetic field is relatively small (<0.1 mV) and can be avoided by eliminating loops or twisting wires.

Some of the basic physics required for this section is summarized in Appendix 1.1. The calculus-challenged reader can consult Appendix 3.2 for the derivatives used in the following examples.

(a) Electromagnetic. In this example, we consider the effect of a magnetic field that is associated with a power line current (I) with an amplitude of 1 A, and line frequency of 60 Hz. Such a current generates a magnetic field (B) at 1 m distance (d) with amplitude (Fig. 3.4A, B):

$$B = \frac{\mu}{2\pi}\frac{I}{d} = 2 10^{-7} \quad \text{T(Tesla)} \tag{3.17}$$

using the magnetic permeability value for vacuum $\mu_0 = 4\pi\, 10^{-7}$.
For a loop enclosing 10^{-2} m^2 and assuming (to simplify the example) that the magnetic field's orientation is perpendicular to the surface area S enclosed by the loop, this translates into a flux:

$$\Phi_B = BS = 2\,10^{-9} \sin(2\pi 60\,t)\,\text{Wb(Weber)}$$

Calculating the amplitude of the potential difference in the loop (V) from the derivative of the flux (Appendices 1.1 and 3.2) generates

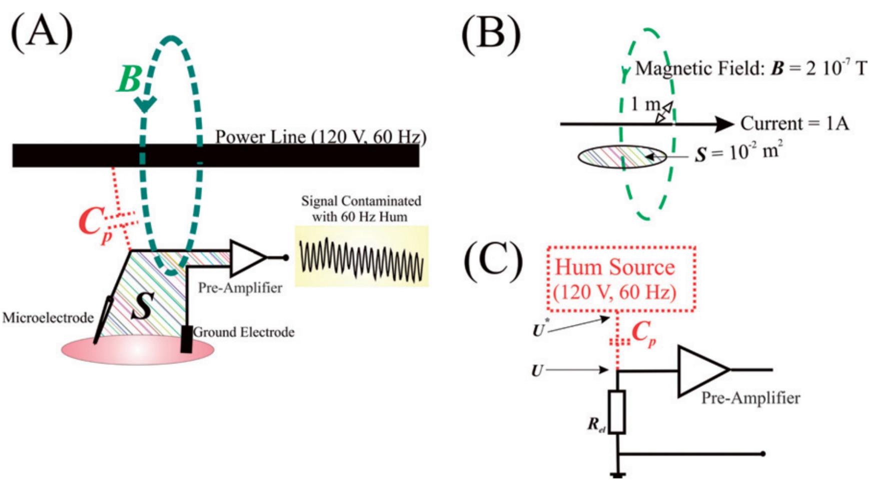

Figure 3.4 Electromagnetic noise caused by a power line can be modeled by the effect of a magnetic flux through the surface S formed between the electrodes and the capacitance C_p between the power line and the input of the preamplifier. (B) Simplified diagram of the magnetic effect in which a magnetic field of 2 10^{-7} T generated by a 1 A current passes through a surface S at 1 m distance. (C) Simplified diagram of the electrostatic effect.

$$V = \frac{d\Phi_B}{dt} = 2\,10^{-9} 2\pi 60 \cos(2\pi 60\, t) \approx \pm 0.75\,\mu\text{V}$$

To calculate the amplitude of the noise in the preceding equation, we only consider the extreme values (±1) of cos(2π 60 t). Thus, the v_{eff} of this sinusoidal signal (Appendix 3.3) is ≈ 0.71 × 0.75 ≈ 0.53 μV.

(b) Electrostatic. The same power line producing the electromagnetic interference characterized in Figure 3.4 also has an electrostatic effect on the input circuitry of the preamplifier. We represent the AC power line as a hum source (U^*, Fig. 3.4C) of 120 V at 60 Hz close to the preamplifier input. The input is also connected to a 10 MΩ (1 MegaOhm = 10^6 Ω) resistance (R_{el}) representing the microelectrode. The conductors of the front end in this setup form a capacitance with conductors that carry the noise signal, the so-called parasitic capacitance. This parasitic capacitance C_p is typically very small, on the order of 10 fF (1 femtoFarad = 10^{-15} F). The current i_c through C_p is the derivative of its charge (Appendix 1.1):

$$i_c = C_p \frac{d(U^* - U)}{dt} \text{with} : U^* = 120 \sin(2\pi 60 t) \tag{3.18}$$

Considering that $U^* \gg U$, we can simplify this to the following approximation:

$$i_c \approx C_p \frac{dU^*}{dt} \tag{3.19}$$

At the level of the preamplifier's input, the effect of current i_c on the input potential is

$$U = i_c R_{el} \approx R_{el} C_p \frac{dU^*}{dt} \tag{3.20}$$

Here we only consider the effect of i_c on the measured potential U. Because we are interested in the noise component, we can ignore any other sources at the preamplifier's input. The derivative (Appendix 3.2) in the preceding expression is

$$2\pi 60 \times 120 \underbrace{\cos(2\pi 60\, t)}_{\pm 1} \approx \pm 4.5 10^4$$

This outcome, multiplied by $R_{el}C_p = 10^{-7}$, results in a noise amplitude of ±4.5 mV. The v_{eff} of this sinusoidal signal (Appendix 3.3) is therefore ≈ 0.71 × 4.5 ≈ 3.2 μV.

As shown in the examples, hum from an electrostatic noise source is usually much larger than the electromagnetic component. This electrostatic noise must be eliminated by shielding or removing the source.

3. In addition to the noise added by passive components such as resistors, active, elements also add noise. Therefore, the application of low-noise amplifiers "early" in the chain (before major amplification steps) is desirable. Typically an active component will add 1-100 μV of noise.
4. The discretization error made at the ADC can also be considered a noise source, the so-called quantization noise. The level of this noise depends on the range and the resolution of the ADC. Assuming an ADC that truncates the sample values, all values in between 0 and 1 become 0, values in between 1 and 2 become 1, and so on. This imprecision is exactly one unit (i.e., the precision) of the analog-to-digital converter, and this applies for the whole range of the converter. The occurrence of truncation errors within the ADC precision can be depicted as a probability density distribution for the added noise. For the sake of this example, let's use an A/D precision of q μV (1 μV = 10^{-6} V); we will obtain a uniform distribution (as in Fig. 3.2B where $q = 6$) if we assume that the signal we sample is equally likely to occur anywhere within each of the units of the ADC. This is a fairly reasonable assumption since we sample a continuous signal and the ADC steps are relatively small. Knowing the PDF of the noise, we can obtain the v_{eff} — that is, the standard deviation of the noise PDF (see Equation (3.14)) — by calculating the square root of $E\{(x - \langle x \rangle)^2\}$ (Equation (3.11)). First we obtain $E\{x\} = \langle x \rangle$ using Equation (3.9):

$$E\{x\} = \langle x \rangle = \int_{-\infty}^{\infty} x\, p(x) dx \tag{3.21}$$

We can change the integration limits from $[-\infty, \infty]$ to $[0, q]$, because outside this domain $p(x) = 0$ and inside $p(x) = 1/q$:

$$\int_{-\infty}^{\infty} x\, p(x) dx = \int_0^q x \underbrace{p(x)}_{\frac{1}{q}} dx \tag{3.22}$$

Because $1/q$ is a constant and we are integrating with respect to x (Appendix 3.2), this expression evaluates to:

$$= \frac{1}{q}\int_0^q x\, dx = \frac{1}{q}\left[\frac{1}{2}x^2\right]_0^q = \frac{q}{2}\mu\text{V} \tag{3.23}$$

Of course, we could have seen by inspection of the example of Figure 3.2B where $q = 6$ that the mean $= q/2 = 3$. Subsequently, we use $\langle x \rangle = q/2$ and $p(x) = 1/q$ between 0 and q in Equation (3.11):

$$E\left\{(x-\langle x\rangle)^2\right\} = \int_{-\infty}^{\infty}\left(x-\underbrace{\langle x\rangle}_{\frac{q}{2}}\right)^2 \underbrace{p(x)}_{\frac{1}{q}}dx = \int_0^q\left(x-\frac{q}{2}\right)^2\frac{1}{q}dx \tag{3.24}$$

Because $1/q$ is a constant and using $(A - B)^2 = A^2 - 2AB + B^2$, we obtain

$$= \frac{1}{q}\int_0^q\left(x^2 - qx + \frac{q^2}{4}\right)dx \tag{3.25}$$

Evaluating the integral (Appendix 3.2):

$$= \frac{1}{q}\left[\frac{1}{3}x^3 - \frac{1}{2}qx^2 + \frac{q^2}{4}x\right]_0^q = \frac{q^2}{12}\mu\mathrm{V}^2 \tag{3.26}$$

The value v_{eff} is then the square root of this variance term — that is, $v_{eff} = \sqrt{\frac{q^2}{12}}\mu\mathrm{V}$.

In state-of-the-art electrophysiology equipment, quantization noise is a few microvolts or less. It is not uncommon to use at least a 12-bit converter. Taking into account amplification of 1000× with an analog ADC input range of 10 V (±5 V), we obtain an analog range of 10/1000 = 0.001 V = 10 mV at the amplifier's input. This results in quantization noise values on the order of microvolt; in this example, $q = \frac{10}{2^{12}} = 0.0024\,\mathrm{mV} = 2.4\ \mu\mathrm{V}$.

In the preceding example, we evaluated the effect of a truncation at the conversion step. If we consider a converter that rounds instead of truncates, the noise characteristics are similar because the PDF (such as the one shown in Fig. 3.2B) only shifts to the left (zero mean). The shape of the distribution, its range, and, consequently, the standard deviation remain unaltered.

From the results shown in these examples, it may be clear that with modern equipment, low noise recordings are indeed feasible. However, often the amplitude of the noise is comparable to the amplitude of different types of biopotentials (Fig. 3.5), indicating that strategies for noise reduction are required. Enemy number 1 in any recording of biopotentials

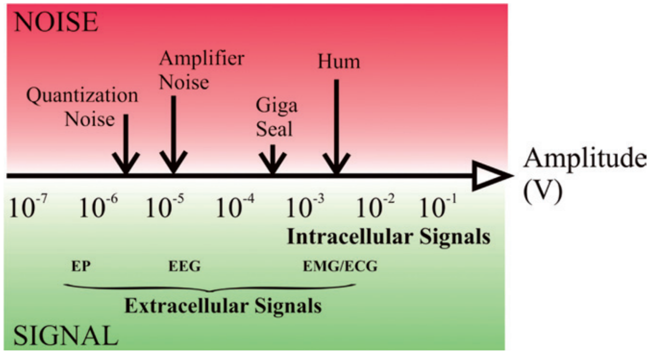

Figure 3.5 Overview of the amplitude of typical biopotentials and different types of noise.

(or low-level transducer signals with similar amplitudes) is hum. As we will see, hum as a nonrandom noise source may even play a role in spoiling signal averaging results.

APPENDIX 3.1

1. Less strict definitions of stationarity and ergodicity exist. A random process with a mean that is time invariant and an autocorrelation function (Chapter 8) that is only dependent on time lag τ is called a ***wide sense stationary*** process. For ergodicity, one may also use more relaxed definitions (e.g., a random stationary process is ***ergodic in the mean*** if at least the mean can be estimated with a time average of a sample function).
2. Because the sample functions from an ergodic process are statistically equivalent, ***an ergodic process is stationary*** and, although there are exceptions, a stationary process will usually also be ergodic. A somewhat trivial example of such an exception is sample function $x(t) = |Y \sin(2\pi(t + Y))|$, in which Y is selected randomly from the same PDF but selection occurs only once for each sample function.
3. A thorough discussion of stationarity and ergodicity is beyond the scope of this text, and for measured time series we will use the labels stationary and ergodic as fancy ways to state optimistically that we believe that our signal at hand allows us to estimate relevant statistics from time averages. In signal processing literature, it is not uncommon to select sample epochs that seem stationary and representative of the signal as a whole, and to use this Gestalt as a reason to declare (explicitly or implicitly) stationarity and ergodicity. Strict tests to

provide proof of these assertions are not available because we are not old enough and do not live long enough to observe a process from $-\infty$ to ∞. Operational definitions for more reasonable time spans do exist, but in practice such tests are rarely used to justify stationarity or ergodicity assumptions.

APPENDIX 3.2

In this book it is assumed that the student is familiar with basic calculus. For refreshing your knowledge of differentiation and integration, see Bo as (1966), Jordan and Smith (1997), or any textbook on these mathematical techniques. This appendix provides a quick reference for those who need a reminder of the most common equations that are used throughout the text.

f (function)	df/dt (derivative)	$\int f\,dt$ (integral)
a (a constant)	0	$ax + C$
x^n for $n \neq -1$	nx^{n-1}	$\frac{1}{n+1}x^{n+1} + C$
x^{-1}	$-1x^{-2}$	$\ln(x) + C$
e^x	e^x	$e^x + C$
$\sin(x)$	$\cos(x)$	$-\cos(x) + C$
$\cos(x)$	$-\sin(x)$	$\sin(x) + C$

Useful rules are

1. The ***chain rule*** is used when differentiating a function $f(u)$ with $u = u(t)$:

$$\frac{df}{dt} = \frac{df}{du}\frac{du}{dt} \tag{A3.2-1}$$

For example the, derivative of $\sin(at)$ with $u = at$ is:

$$\left.\begin{aligned} \frac{df}{du} &= \frac{d\sin(u)}{du} = \cos(u) = \cos(at) \\ \frac{du}{dt} &= \frac{d(at)}{dt} = a \end{aligned}\right\} \rightarrow \frac{d[\sin(at)]}{at} = \frac{df}{du}\frac{du}{dt} = a\cos(at)$$

2. ***Differentiation and integration by parts*** for function $f = uv$.

a. ***Differentiation*** (here we use the notation f' for the derivative):

$$\frac{df}{dt} = f' = (uv)' = uv' + u'v \tag{A3.2-2}$$

For instance, differentiate $f = 3xe^{ax}$. Using this approach, we have the following:

$$u = 3x \quad \text{and} \quad v = e^{ax}$$

Getting the differentials required:

$$u' = 3 \quad \text{and} \quad v' = ae^{ax}$$

Substituting this into Equation (A3.2-2), we obtain the solution for the differential:

$$f' = uv' + u'v = 3axe^{ax} + 3e^{ax} = 3e^{ax}(1 + ax)$$

b. ***Integration:***

$$\int u dv = uv - \int v du \tag{A3.2-3}$$

We integrate the same function as in Section 2a: $f = 3xe^{ax}$.
Using integration by parts, we have

$$u = 3x \quad \text{and} \quad dv = e^{ax}dx$$

Getting the other expressions required:

$$du = 3dx \quad \text{and} \quad v = \frac{1}{a}e^{ax}$$

Substituting this into Equation (A3.2-3):

$$\begin{aligned}\int 3xe^{ax}dx &= uv - \int v du = 3x\frac{1}{a}e^{ax} - \int \frac{1}{a}e^{ax}3dx \\ &= \frac{3}{a}xe^{ax} - \frac{3}{a}\int e^{ax}dx = \frac{3}{a}xe^{ax} - \frac{3}{a}\left[\frac{1}{a}e^{ax}\right] + C = \frac{3}{a}e^{ax}\left[x - \frac{1}{a}\right] + C\end{aligned}$$

As we can see, this approach works well in this example because the evaluation of $\int v du$ is easier than the integral of $\int u dv$.

APPENDIX 3.3

The v_{eff} of a sinusoidal signal with amplitude A can be calculated with Equation (3.14). Consider a sine wave that fluctuates around zero (mean = 0) with frequency f (= $1/T$) for n periods (i.e., a time interval equal to nT):

$$v_{eff}^2 = \frac{1}{nT}\int_0^{nT} A^2 \sin^2(2\pi ft)\,dt \tag{A3.3-1}$$

Taking the constant A^2 out of the integration and using the trigonometric equality:

$$\sin^2\left(\frac{1}{2}\alpha\right) = \frac{1}{2} - \frac{1}{2}\cos(\alpha)$$

we obtain:

$$v_{eff}^2 = \frac{A^2}{nT}\int_0^{nT}\left(\frac{1}{2} - \frac{1}{2}\cos(4\pi ft)\right)dt \tag{A3.3-2}$$

Separating the terms in the integral:

$$v_{eff}^2 = \frac{A^2}{nT}\underbrace{\int_0^{nT}\left(\frac{1}{2}\right)dt}_{\left[\frac{t}{2}\right]_0^{nT}} - \frac{A^2}{nT}\underbrace{\int_0^{nT}\frac{1}{2}\cos(4\pi ft)\,dt}_{0} \tag{A3.3-3}$$

Evaluation of the first term in Equation (A3.3-3) is the integration of a constant (1/2). Because the second term in Equation (A3.3-3) is the integral of a cosine function over an integer number of periods, the net area enclosed by the wave is zero and therefore this integral evaluates to zero. Equation (A3.3-3) evaluates to

$$v_{eff}^2 = \frac{A^2}{nT}\frac{nT}{2} = \frac{A^2}{2} \tag{A3.3-4}$$

The v_{eff} of a sine wave over a full number of periods is therefore equal to

$$\sqrt{\frac{A^2}{2}} = \frac{A}{2}\sqrt{2} \approx 0.71A$$

APPENDIX 3.4

In this appendix we explore the use of Laplace and Fourier transforms to facilitate the determination of parameters that are associated with probability density functions (PDFs). If you are not yet familiar with the Laplace and Fourier transforms, you can skip this part and address it later. The Laplace transform is frequently used in statistics to characterize combined processes with different probability density distributions or to generate the moments of a PDF.

If T is a non-negative random variable drawn from a PDF $f(t)$ with moments $E(T)$, $E(T^2)$, . . . defined as

$$E(T^n) = \int_0^\infty t^n f(t)\,dt \tag{A3.4-1}$$

Note that the integration is from $0 \to \infty$, because T is non-negative — that is, $f(t) = 0$ for $t < 0$. The Laplace transform of $f(t)$ is

$$F(s) = \int_0^\infty f(t) e^{-st} dt = E(e^{-sT}) \tag{A3.4-2}$$

The exponential can be written as a series:

$$\begin{aligned} F(s) &= E\left(1 - \frac{sT}{1!} + \frac{s^2T^2}{2!} - \frac{s^3T^3}{3!} + \frac{s^4T^4}{2!} \ldots\right) \\ F(s) &= E\left(\sum_{k=0}^{\infty} (-1)^k \frac{s^kT^k}{k!}\right) = \sum_{k=0}^{\infty} (-1)^k \frac{s^k}{k!} E(T^k) \end{aligned} \tag{A3.4-3}$$

As the last expression shows, the Laplace transform generates the moments $E(T^k)$. Sometimes it is easier to use this property to find the moments of a distribution than to explicitly evaluate the integral in Equation (A3.4-1), for instance, in the exponential distribution associated with a Poisson process (not to be confused with a Poisson distribution, see Chapter 14). The Poisson process $f(t) = \rho\, e^{-\rho t}$ has a Laplace transform $F(s) = \rho/(\rho + s)$. This Laplace transform can be presented as an infinite series:

$$F(s) = \frac{\rho}{\rho + s} = \sum_{k=0}^{\infty} (-1)^k \frac{s^k}{\rho^k} \tag{A3.4-4}$$

Comparing this series with the generic one, we can establish that for the Poisson process: $E(T^k) = k!/\rho^k$, indicating that the mean $E(T) = 1/\rho$ and the variance $\sigma^2 = E(T^2) - E(T)^2 = 2/\rho^2 - (1/\rho)^2 = 1/\rho^2$. Thus, for the Poisson process PDF, the mean and standard deviations are both equal to $1/\rho$.

A transform very similar to the Fourier transform of a PDF is also used for studying the propagation of noise through a system with a known transfer function. This transform of the PDF is called the *characteristic function* $\psi(\omega) = \int_{-\infty}^{\infty} f(t) e^{j\omega t} dt$ of the PDF, where the only difference from the Fourier transform is the sign of ω. The characteristic function can also be used to determine the moments of the PDF by taking the derivatives of $\psi(\omega)$: $\frac{d^n \psi(\omega)}{d\omega^n} = \int_{-\infty}^{\infty} j^n t^n f(t) e^{j\omega t} dt$. For $\omega = 0$:

$$\left[\int_{-\infty}^{\infty} j^n t^n f(t) e^{j\omega t} dt \right]_{\omega=0} = \int_{-\infty}^{\infty} j^n t^n f(t) dt = j^n E(T^n) \tag{A3.4-5}$$

This equation can sometimes make it easier to determine the moments of a distribution from a table of Fourier transforms. Again, we can use the example of the Poisson process $f(t) = \rho\, e^{-\rho t}$ with its characteristic function: $\rho/(\rho - j\omega)$. Note that we used the Fourier transform of the PDF and simply changed the sign of ω. To establish the first moment $n = 1$; the first derivative of this characteristic function is as follows (remember that the derivative of a quotient u/v is $u'v - uv'/v^2$; here $u = \rho$ and $v = \rho - j\omega$):

$$\frac{d\psi(\omega)}{d\omega} = \left[\frac{-\rho(-j)}{(\rho - j\omega)^2} \right]_{\omega=0} = \frac{\rho j}{\rho^2} = \frac{j}{\rho} \tag{A3.4-6}$$

According to Equation (A3.4-5), this expression equals $jE(T) \rightarrow$ the expected mean value $E(T) = 1/\rho$.