includes active dendritic spines (Shepherd et al. 1985). It has been postulated that enhanced propagation of voltage changes induced by synaptic input can occur because spine–spine excitations occur in a way similar to the saltatory conduction of spikes in a myelinated nerve.

Propagation in nerve bundles

When several nerve axons form a nerve tract, there is the possibility of electrical interactions between individual fibers through the extracellular space. This was demonstrated by Katz and Schmidt (1940) who found that when a pulse in one fiber was followed by a faster pulse in an adjacent fiber, the second pulse would catch up with the first and the two pulses would travel together in a locked-in fashion.

There have been a few theoretical investigations of such interactions. Scott and Luzader (1979) and Eilbeck, Luzader, and Scott (1981) employed the coupled Fitzhugh–Nagumo-type equations

$$v_{1,t} = (1-\alpha)v_{1,xx} - \alpha v_{2,xx} + f(v_1) - w_1, \tag{8.217}$$

$$w_{1,t} = b(v_1 - \gamma w_1), \tag{8.218}$$

$$v_{2,t} = (1-\alpha)v_{2,xx} - \alpha v_{1,xx} + f(v_2) - w_2, \tag{8.219}$$

$$w_{2,t} = b(v_2 - \gamma w_2), \tag{8.220}$$

where $\alpha \ll 1$ is a coupling parameter. The solutions of the traveling-wave equations were expanded as a perturbation series in powers of α and an expression obtained for the first-order term in the correction to the speed of propagation when $f(\cdot)$ was piecewise linear.

Bell (1981) has considered a pair of coupled Hodgkin–Huxley axons and provided conditions for the existence of a traveling-pulse solution on both fibers. A result was also obtained in the case of $n > 2$ fibers and for other dynamical equations representing the behavior of active membrane (Bell and Cook 1978, 1979).

9

The stochastic activity of neurons

9.1 Introduction

All of the models of nerve-cell activity we have considered thus far have been *deterministic*. They have consisted of a differential equation (with a threshold condition imposed) or a system of differential equations in conjunction with a given input current and given boundary–initial-value conditions.

To illustrate, consider the linear Lapicque model of Chapter 3. If $V(t)$ is the depolarization at time t, C is the neuron capacitance, R is the neuron resistance, and $I(t)$ is the input current, then

$$C\,dV/dt + V/R = I(t), \qquad t > 0,\ V < \theta,\ V(0) = 0, \tag{9.1}$$

with a spike generated when V reaches or exceeds θ. If $I(t)$ is constant, then the predicted time between action potentials is always the same and is given by (3.52). If $I(t) = I_0 + I_1\cos(\omega t + \phi)$, representing a cyclic input, the time interval between spikes is *variable* (see Figure 3.21) but *the sequence of times of occurrence of spikes is completely and uniquely determined*.

Such deterministic models are inadequate for the description of the activity of real neurons. Assuming for now that the mathematical model (differential equations, boundary conditions, threshold condition, if necessary) is valid, the input current is rarely, if ever, known with certainty. This is true even in controlled experiments such as the current-injection experiments described in Chapter 1 and analyzed in Chapter 8. Even in those experiments the time between action potentials is variable.

In most experimental studies of the activity of nerve cells in more natural conditions, very little, if anything, is known about the inputs to the cell under study. Indeed, one of the challenges has been to ascertain the nature of the input even in relatively simple situations

(see Section 6.5). Certain sensory inputs may be controlled but still the actual inputs to the nerve cells are not known. The observed quantity is the train of spikes, or possibly the time course, of the intracellular or extracellular potential, not the input currents, and so forth.

Although action potentials are not instantaneous, it is customary to assign them occurrence times, which may, for example, be the time at which the voltage apparently attains threshold.

Definition

Let $\{\Theta_k, k = 0, 1, 2, \ldots\}$ be a sequence of times at which a nerve cell emits action potentials, with $\Theta_0 = 0$ and $\Theta_0 < \Theta_1 < \Theta_2 < \cdots$. The kth *interspike interval* (ISI) is

$$T_k = \Theta_k - \Theta_{k-1}, \qquad k = 1, 2, \ldots. \tag{9.2}$$

The first measurements that revealed the variability of the ISI were on the muscle spindles of frogs (Brink, Bronk, and Larrabee 1946; Buller, Nicholls, and Strom 1953; Hagiwara 1954). Under conditions of constant tension of the muscle, the ISIs were quite variable and this variability was greatest when the muscle was in its unstretched state. Some results from one such experiment are shown in Figure 9.1, where t_k, the value of T_k, is plotted against k. It can be seen that there is a tendency for the interspike intervals to become longer and longer (*adaptation*), but there is also a haphazard fluctuation.

Spontaneous activity

The term *spontaneous activity* with reference to neural activity is not well defined. It may be loosely described as the activity of a

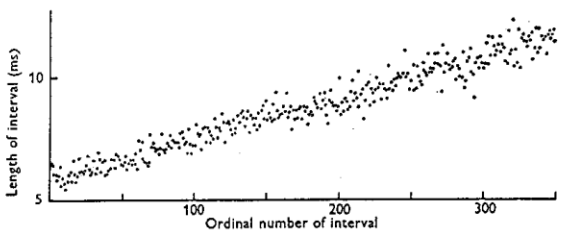

Figure 9.1. Length of interspike intervals versus order of appearance for a muscle spindle held at fixed tension. [From Buller et al. (1953). Reproduced with the permission of The Physiological Society and the authors.]

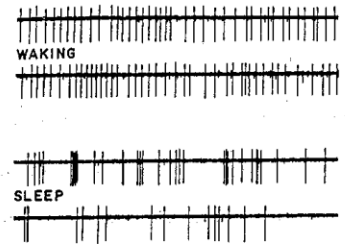

Figure 9.2. Spike trains from a pyramidal tract cell of a monkey while the animal was awake (upper two traces) and asleep (lower two traces). [From Evarts (1964). Reproduced with the permission of The American Physiological Society and the author.]

neuron, with focus on the collection of action potentials it emits, in the absence of an intended input or when there is no obvious input from other cells. The majority of, but not all, CNS neurons exhibit spontaneous activity consisting of an irregular train of impulses. Examples are shown in Figure 9.2, where the spiking activity of a pyramidal tract cell of a monkey is shown during periods when the animal was awake and asleep. It can be seen that the description of such spike trains, short of enumerating all the occurrence times of impulses, is difficult. Usually the collection of interspike intervals is collected into a histogram; examples from another experiment are shown in Figure 9.3.

Given that a sequence of occurrence times of action potentials characterizes the activity of a nerve cell, it is then of interest to see how the train of impulses changes under various conditions. Since the spike trains themselves and the dynamical processes that cause them are usually random, it is clear that models of the activity of real neurons must be *stochastic* and not deterministic. The introduction of randomness makes the mathematical theory more difficult and progress has been made only with what are physically very simple models. Previous works dealing with the stochastic aspects of neural activity are those of Goel and Richter-Dyn (1974), Holden (1976b), MacGregor and Lewis (1977), Ricciardi (1977), and Sampath and Srinivasan (1977).

There have been numerous experimental studies of the stochastic activity of nerve cells. We will mention some representative ones but do not have space to describe the results. In the investigations of the following list, the focus is sometimes on the spontaneous activity or

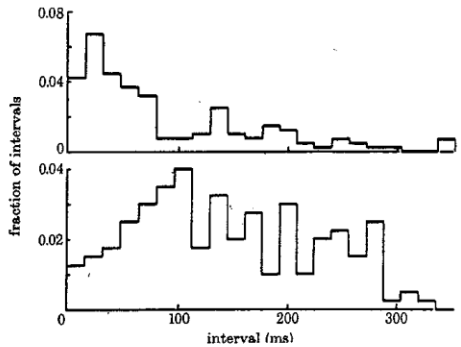

Figure 9.3. Histograms of ISIs for a cell in the cat visual cortex when the animal was "sitting in complete darkness awake, but relaxed" (upper figure) and when it was "alarmed by a hissing noise." [From Burns and Webb (1976). Reproduced with the permission of The Royal Society and the authors.]

on the effects of various influences such as state of arousal, learning, drugs, sensory inputs, epileptogenesis, radiation, and so forth. For studies on *muscle spindles*, in addition to the references already given, see Stein and Matthews (1965); on cells of the *auditory system*, see Gerstein and Kiang (1960), Grossman and Viernstein (1961), Rodieck, Kiang, and Gerstein (1962), Goldberg, Adrian, and Smith (1964), Pfeiffer and Kiang (1965), Gerstein, Butler, and Erulkar (1968), Molnar and Pfeiffer (1968), Kiang and Moxon (1974), Johnson and Kiang (1976), and Ryan and Miller (1977); on cells in the *visual pathway*, see Levick et al. (1961), Bishop, Levick, and Williams (1964), and Burke and Sefton (1966a, b); on cells in the *cerebral cortex* including pyramidal cells, see Evarts (1964), Koike et al. (1970), Whitsel, Roppolo, and Werner (1972), O'Brien, Packham, and Brunnhoelzl (1973), Steriade, Wyzinski, and Apostol (1973), Wyler and Fetz (1974), Burns and Webb (1976), Webb (1976a, b), Whitsel, Schreiner, and Essick (1977), and Schreiner, Essick, and Whitsel (1978); on *hippocampal pyramids*, see Bassant (1976); on *chemoreceptors*, see Silk and Stein (1966); on cells of the *dorsalspinocerebellar tract*, see Jansen, Nicolaysen, and Rudjord (1966) and Pyatigorskii (1966); on *spinal motoneurons*, see Calvin and Stevens (1968); on cells of the *thalamus*, see Poggio and Viernstein (1964), Nakahama et al. (1966), Baker (1971), Lamarre, Filion, and Cordeau (1971), Dormont (1972), and Benoit and Chataigner (1973); on cells of the *amygdaloid complex*, see Eidelberg, Goldstein, and Deza (1967); on *Purkinje cells of the cerebellum*, see Braitenberg et al. (1965), Pellet et al. (1974), and Woodward, Hoffer, and Altman (1974); on cells in certain *invertebrates*, see Firth (1966), Junge and Moore (1966), Buno, Fuentes, and Segundo (1978), and Holden and Ramadan (1979); on cells of the *vestibular system*, see Wylie and Pelpel (1971), Goldberg, Fernandez, and Smith (1982), and Goldberg, Smith, and Fernandez (1984); and on cells of the *reticular formation*, see Sparks and Travis (1968) and Syka, Popelar, and Radil-Weiss (1977). In addition, in some experiments nerve cells have been deliberately subjected to random stimulation (Wilson and Wyman 1965; Redman and Lampard 1967, 1968; Redman, Lampard, and Annal 1968; Lampard and Redman 1969; Guttman, Feldman, and Lecar 1974; Bryant and Segundo 1976).

Synaptic noise

Noise has also been observed in other nerve-cell preparations. Fatt and Katz (1952) observed randomly occurring synaptic potentials called miniature endplate potentials (m.e.p.p.'s) at the frog neuromuscular function. Their amplitudes were on the order of 0.5 mV and their frequencies were between 0.1 to 100 impulses/s for various muscle fibers. These spontaneous potentials have been important in helping elucidate the nature of synaptic transmission and are studied in Sections 9.3 and 9.4. Spontaneous miniature EPSP's and IPSP's have also been observed in spinal motoneurons of the frog (Katz and Miledi 1963) and the cat (Calvin and Stevens 1968) as well as in pyramidal cells of the cat motor cortex (Watanabe and Creutzfeldt 1966). Sometimes these spontaneous postsynaptic potentials are referred to as *synaptic noise*.

Membrane noise

Even when apparently steady conditions prevail, there are observed small fluctuations in the electrical potential across the nerve-cell membrane. These fluctuations have been attributed to the back-and-forth motion of ions and electrons due to thermal agitation (*Brownian motion*), the discrete nature of currents through the membrane (*shot noise*), and conductance changes induced by the random opening and closing of ion channels. These random fluctuations in membrane potential may be collectively referred to as *membrane noise*.

9.2 Probability and random variables

Probability enters modeling when we do not know, or cannot prescribe with certainty, the conditions or mechanisms that prevail in the system of interest. For example, it is not possible to predict the times of occurrence of the action potentials in Figure 9.2. On the basis of an ISI histogram all we can do, on the basis of collected data, is make statements of the kind, "there is a 40% chance that a spike will occur in the next 10 ms." Statements such as this can be collected into more precise statements in terms of *random variables* and *random processes*. We will give a brief review of some of the basic concepts – texts such as Feller (1968) and Chung (1979) should be consulted for thorough introductory treatments. A wide range of applications is considered in Tuckwell (1988a).

When an experiment is performed whose outcome is uncertain, the collection of possible elementary outcomes is called a *sample space*. Roughly speaking, a *random variable* is an observable that takes on numerical values with certain probabilities. It is a real-valued function defined on the elements of a sample space. We adopt the convention that random variables themselves are denoted by capital letters, whereas the values they take on are denoted by lowercase letters.

Discrete random variables take on finitely many or countably infinitely many values. The most important such random variables for us are binomial and Poisson.

A *binomial* random variable X has the probability law

$$p_k = \Pr\{X = k\} = \binom{n}{k} p^k q^{n-k}, \qquad k = 0, 1, 2, \ldots, n, \tag{9.3}$$

where $0 \le p \le 1$, $q = 1 - p$, and n is a positive integer. The number of heads in n tosses of a coin that has probability p of landing heads has this probability law.

A *Poisson* random variable with parameter $\lambda > 0$ takes on non-negative integer values and has the probability law

$$p_k = \Pr\{X = k\} = \frac{e^{-\lambda}\lambda^k}{k!}, \qquad k = 0, 1, 2, \ldots. \tag{9.4}$$

For both of these random variables the total probability mass is unity

$$\sum_k p_k = 1. \tag{9.5}$$

Continuous random variables take on a continuum of values. For any random variable the *distribution function* is

$$F(x) = \Pr\{X \le x\}. \tag{9.6}$$

Usually the probability law of a continuous random variable can be expressed through its *probability density* $p(x)$, which is the derivative of its distribution function. Then, roughly speaking, $p(x)\,dx$ is the probability that the random variable takes on values in $(x, x + dx]$. In what follows two kinds of continuous random variables are important. One is the *normal*, or *Gaussian*, random variable with density

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \qquad -\infty < x < \infty, \tag{9.7}$$

where μ and σ^2 are constants. The other has a *gamma* density

$$p(x) = (\lambda/\Gamma(r))(\lambda x)^{r-1} e^{-\lambda x}, \qquad x > 0, \tag{9.8}$$

where $\lambda > 0$ and $r > 0$ are constants and $\Gamma(r) = \int_0^\infty e^{-x} x^{r-1}\,dx$ is the gamma function. A special case is that of an *exponentially distributed* random variable with $r = 1$ so the density is $p(x) = \lambda e^{-\lambda x}$. For continuous random variables the total probability is also unity

$$\int p(x)\,dx = 1, \tag{9.9}$$

where the range of integration is $(-\infty, \infty)$ for a normal random variable and $(0, \infty)$ for a gamma variate.

Mean, variance, and covariance

The *mean*, or *expected value*, of an integer-valued random variable is

$$E[X] = \sum_k k p_k. \tag{9.10}$$

This gives $E[X] = np$ for the binomial and $E[X] = \lambda$ for the Poisson random variable.

For continuous random variables,

$$E[X] = \int x p(x)\,dx. \tag{9.11}$$

This gives $E[X] = \mu$ for the normal and $E[X] = r/\lambda$ for the gamma random variable. Note that expectation is a linear operation, so $E[aX + bY] = aE[X] + bE[Y]$.

The *second moment* is $E[X^2]$ and the *variance*, which tells us how much scatter there is about the mean, is

$$\mathrm{Var}[X] = E\left[(X - E[X])^2\right] = E[X^2] - E^2[X]. \tag{9.12}$$

The variances of the random variables considered above are: binomial, npq; Poisson, λ; normal, σ^2; and gamma, r/λ^2.

Conditional probability and independence

Let A and B be two random events. The *conditional probability* of A given B is defined as

$$\Pr\{A|B\} = \Pr\{AB\}/\Pr\{B\}, \tag{9.13}$$

where AB means both A and B occur and it is assumed that $\Pr\{B\} \neq 0$. That is, only those occurrences of A that are simultaneous with those of B are taken into account. This extends to random variables. For example, if X and Y are two random variables, defined on the same sample space, taking on values x_i, $i = 1, 2$, and y_j, $j = 1, 2, \ldots$, respectively, then the conditional probability that $X = x_i$ given $Y = y_j$ is

$$\Pr\{X = x_i | Y = y_j\} = \Pr\{X = x_i, Y = y_j\}/\Pr\{Y = y_j\}. \tag{9.14}$$

Note that (9.13) rearranges to $\Pr\{AB\} = \Pr\{B\}\Pr\{A|B\}$.

The *conditional expectation* of X given $Y = y_j$ is

$$E[X|Y = y_j] = \sum_i x_i \Pr\{X = x_i | Y = y_j\}. \tag{9.14A}$$

The expected value of XY is

$$E[XY] = \sum_{i,j} x_i y_j \Pr\{X = x_i, Y = y_j\}, \tag{9.15}$$

and the *covariance* of two random variables X and Y is

$$\mathrm{Cov}[X, Y] = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y]. \tag{9.16}$$

The covariance is a measure of the linear dependence of X and Y.

If X and Y are *independent*, the value of Y should have no influence on the probability that X takes on its values. Hence we may define X and Y as independent if

$$\Pr\{X = x_i | Y = y_j\} = \Pr\{X = x_i\}, \tag{9.17}$$

for all i, j. Equivalently, $\Pr\{X = x_i, Y = y_j\} = \Pr\{X = x_i\}\Pr\{Y = y_j\}$, which leads, in the case of independent random variables, to

$$E[XY] = E[X]E[Y]. \tag{9.17A}$$

It also follows that if X and Y are independent, $\mathrm{Cov}[X, Y] = 0$ (but note that $\mathrm{Cov}[X, Y] = 0$ does not always imply that X and Y are independent). If X_i, $i = 1, 2, \ldots, n$, are mutually independent, then

$$\mathrm{Var}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathrm{Var}[X_i]. \tag{9.17B}$$

If A_i, $i = 1, 2, \ldots, n$, are mutually exclusive events and at least one must occur, and B is another event, then

$$\Pr\{B\} = \sum_{i=1}^{n} \Pr\{A_i\}\Pr\{B|A_i\}, \tag{9.17C}$$

which is the *law of total probability*.

Finally, the *characteristic function* of a random variable X is defined as the expectation of the complex random variable $\exp(iuX)$, where u varies from $-\infty$ to $+\infty$:

$$\phi_X(u) = E[e^{iuX}], \qquad u \in (-\infty, \infty). \tag{9.18}$$

(Here $i = \sqrt{-1}$.) Characteristic functions are important because there is a one-to-one correspondence between them and distribution functions.

It is left as an exercise to show that if X is Poisson with parameter λ, then

$$\phi_X(u) = \exp[\lambda(e^{iu} - 1)], \tag{9.18A}$$

whereas if X is normal with mean μ and variance σ^2,

$$\phi_X(u) = \exp[i\mu u - \tfrac{1}{2}u^2\sigma^2]. \tag{9.18B}$$

9.3 The quantum hypothesis in synaptic transmission

In this section we will obtain a probabilistic description of the small voltage changes that occur at spontaneously active synapses. This will be done in the context of nerve–muscle synapses but the same kind of stochastic model should apply at synapses within the nervous system.

In Section 9.1 we mentioned Fatt and Katz's (1952) discovery of randomly occurring miniature endplate potentials at the frog neuromuscular junction. An example of the records they obtained is shown in Figure 9.4. The mean amplitude of the m.e.p.p.'s was about 0.5 mV. This should be compared with the normal postsynaptic response, the endplate potential (e.p.p.), which results when a nerve impulse invades the presynaptic terminal. The e.p.p. has an amplitude between 50 and 70 mV (Kuffler and Nicholls 1976).

In a reduced Ca^{2+} bathing solution the amplitude of the e.p.p. is reduced. Fatt and Katz observed that in such low Ca^{2+} solutions the amplitudes of the reduced e.p.p.'s were approximately in multiples of the amplitudes of the spontaneous m.e.p.p.'s. The *quantum hypothesis* (del Castillo and Katz 1954) was advanced that the normal e.p.p. is the result of the almost simultaneous occurrence of several m.e.p.p.'s.

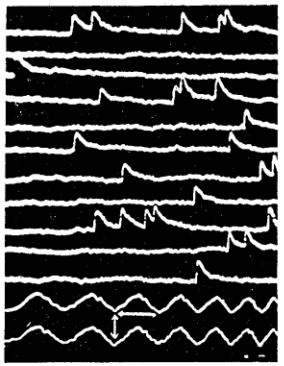

Figure 9.4. Randomly occurring postsynaptic potentials (m.e.p.p.'s) recorded intracellularly from frog muscle fiber. Several records are shown from the same fiber. [From Fatt and Katz (1952). Reproduced with the permission of The Physiological Society and the authors.]

It was speculated that the transmitter (acetylcholine) was, in fact, released in packets containing several thousand molecules corresponding to the quantal EPSP's mentioned in Section 1.5. Furthermore, it seemed likely that the *synaptic vesicles* of diameter about 500 Å, seen with the electron microscope to reside in the presynaptic terminal, were, in fact, the packets of transmitter.

Probability model

We assume there are n sites within the nerve terminal at which transmitter release may occur. When an action potential invades the terminal, release occurs at a random number M of sites. The amplitude of the response due to release at each site is random, that at the ith active site being V_i. Each of the V_i's is assumed to have the same probability distribution and the sites act independently of each other. The amplitude of the e.p.p. is then

$$V = V_1 + V_2 + \cdots + V_M. \tag{9.19}$$

This is a sum in which the number of terms is random.

In the first instance we would assume that M is binomial with parameters n and p, where p is the probability that a site is active. However, if p is small we may use the *Poisson approximation* to the

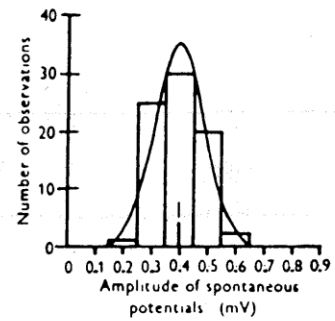

Figure 9.5. Histogram of the amplitudes of the spontaneous m.e.p.p.'s. The smooth curve is a normal density. [From Boyd and Martin (1956). Reproduced with the permission of The Physiological Society and the authors.]

binomial, which has the advantage that now the distribution of M contains only one parameter, $\lambda = np$, the mean number of active sites. The value of λ can be estimated from the fraction of trials, which result in no response, that is,

$$\Pr\{M = 0\} = e^{-\lambda}. \tag{9.20}$$

When M has a Poisson distribution, the random sum V is said to have a *compound Poisson distribution*.

Figure 9.5 shows the histogram of amplitudes of the miniature spontaneous potentials in one preparation. The smooth curve through the histogram is a normal probability density with the same mean and variance. Thus we assume that the V_i's are normal with mean μ and variance σ^2. We are now able to find the density of the amplitude of the e.p.p. V.

The *law of total probability* gives

$$\Pr\{v < V < v + dv\} = \sum_{m=0}^{\infty} \Pr\{v < V < v + dv \mid M = m\} \Pr\{M = m\} \tag{9.21}$$

The conditional probability that $V \in (v, v + dv)$ given $M = m$ is found as follows. If $M = m > 0$, then there are exactly m terms in the sum (9.19)

$$V = V_1 + V_2 + \cdots + V_m. \tag{9.22}$$

These terms are *independent* and each is *normally distributed*. A basic theorem on the sum of independent normal random variables (Parzen 1962, page 17) tells us that V is normally distributed with mean

$$E[V|M=m] = mE[V_1] = m\mu, \tag{9.23}$$

and variance

$$\mathrm{Var}[V|M=m] = m\,\mathrm{Var}[V_1] = m\sigma^2. \tag{9.24}$$

In the case $m=0$ the amplitude has a density, which is a delta function concentrated at $v=0$ with weight $e^{-\lambda}$.

Putting all this together, we arrive at the probability density of the amplitude of the e.p.p.,

$$p_V(v) = \exp(-\lambda)\left\{\delta(v) + \frac{1}{\sqrt{2\pi\sigma^2}}\sum_{m=1}^{\infty}\frac{\lambda^m}{m!\sqrt{m}}\exp\left[\frac{-(v-m\mu)^2}{2m\sigma^2}\right]\right\}, \tag{9.25}$$

which agrees with the formula of Bennett and Florin (1974). It is left as an exercise to show that

$$E[V] = \lambda\mu, \tag{9.25A}$$

$$\mathrm{Var}[V] = \lambda(\mu^2+\sigma^2). \tag{9.25B}$$

Figure 9.6 shows the histogram of amplitudes of the endplate potentials along with the density predicted by (9.25). It can be seen that there is good agreement between the experimental and theoretical result, thus substantiating the quantum hypothesis. Note that since the normal e.p.p. is about 50 mV, then as many as 100 or more quanta are released when the action potential invades the terminal. Supporting evidence for the quantum hypothesis has been obtained in other preparations [see Martin (1977)]. However, a gamma, rather than a normal, density has sometimes been needed for the unit response [see, for example, Bornstein (1978)] and a binomial analysis has sometimes been required when the Poisson approximation has proven inadequate (Miyamoto 1975; Volle and Branisteanu 1976). Brown, Perkel, and Feldman (1976) showed, with the aid of computer simulation, that spatial and temporal nonuniformities in the release parameters can easily lead to erroneous estimates of these quantities.

9.4 The Poisson process

A *random process* is a family of random variables. We will only be concerned with families of random variables parameterized by a continuous index t, the time. The chief physical random processes

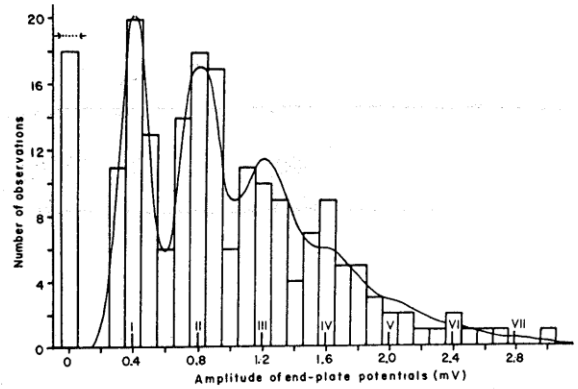

Figure 9.6. Histogram of amplitudes of the e.p.p. The smooth curve is the density of a compound Poisson random variable as given by (9.25). [From Boyd and Martin 1956). Reproduced with the permission of The Physiological Society and the authors.]

with which we will be concerned are the nerve membrane potential and the number of synaptic inputs.

A random variable has a fixed probability law, which is easy to describe, but the probability laws of continuous-time random processes are complicated. If we let $X(t)$ denote the value of a general random process at time t, the whole process is the family $\{X(t), t \geq 0\}$, which we sometimes abbreviate to X.

The first random process we examine is the *simple Poisson process*. This is important for several reasons:

(i) it is a basic, much studied process and can be used as a standard against which some physical processes can be compared;
(ii) it may be used to synthesize more complicated processes of interest;
(iii) it forms a useful approximation for the inputs to a cell when these are many and unsynchronized.

There are many equivalent definitions of the Poisson process [see, for example, Parzen (1962)] one of which is as follows.

Definition

$\{N(t),\ t \le 0\}$ is a simple Poisson process with *intensity* or *mean rate* λ if:

(a) $N(0) = 0$;

(b) given any $0 = t_0 < t_1 < t_2 < \cdots < t_{n-1} < t_n$, the random variables $N(t_k) - N(t_{k-1})$, $k = 1, 2, \ldots, n$, are mutually independent; and

(c) for any $0 \le t_1 < t_2$, $N(t_2) - N(t_1)$ is a Poisson random variable with probability distribution

$$\Pr\{N(t_2) - N(t_1) = k\} = \frac{(\lambda(t_2 - t_1))^k \exp(-\lambda(t_2 - t_1))}{k!}, \qquad k = 0, 1, 2, \ldots. \tag{9.26}$$

Property (a) is just a starting condition. Property (b) puts the Poisson process in the class of processes with *independent increments*. Property (c) tells us that the increments are *stationary* (since only time differences matter) with Poisson distributions. The meaning of λ will become clear shortly.

From (9.26) with $t_1 = 0$ and $t_2 = t$, we see that $N(t)$ is a Poisson random variable with mean and variance equal to λt. Also, with $t_1 = t$ and $t_2 = t + \Delta t$, we find

$$\Pr\{N(t + \Delta t) - N(t) = k\} = \frac{(\lambda\,\Delta t)^k \exp(-\lambda\,\Delta t)}{k!} = \begin{cases} 1 - \lambda\,\Delta t + o(\Delta t), & k = 0, \\ \lambda\,\Delta t + o(\Delta t), & k = 1, \\ o(\Delta t), & k \ge 2, \end{cases} \tag{9.27}$$

where $o(\Delta t)$ means terms that, as $\Delta t \to 0$, approach zero faster than Δt itself. Hence, in very small time intervals, the process is most likely to stay unchanged ($k = 0$) or undergo a step increase of unity ($k = 1$). The value of $N(t)$ will be the number of unit step changes that have occurred in $(0, t]$. A typical *realization* (*sample path*, *trajectory*) of the process will appear as sketched in Figure 9.7A.

In Figure 9.7B we have inserted a cross on the t-axis at the times when $N(t)$ jumps by unity. Each cross may be associated with the

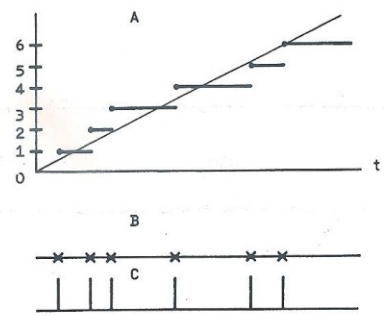

Figure 9.7. A – Realization of a Poisson process. B – Associated realization of the Poisson point process. C – Hypothetical spike train.

occurrence of a certain kind of event such as a postsynaptic potential or an action potential. The crosses (points) form a realization of a *Poisson point process* and $N(t)$ thus records or counts the number of events in $(0, t]$. Figure 9.7C shows a hypothetical spike train, which can be associated with the point process.

The waiting time to an event

Consider any $s > 0$ and let T_1 be the time to the first event occurring after s. Then we find that T_1 *is exponentially distributed with mean* $1/\lambda$.

Proof. The probability that one has to wait longer than t for the first event is the probability that there are no events in $(s, s + t]$. Thus

$$\Pr\{T_1 > t\} = \Pr\{N(s + t) - N(s) = 0\} = e^{-\lambda t}, \qquad t > 0. \tag{9.28}$$

Thus the distribution function of T_1 is $1 - e^{-\lambda t}$ and hence the probability density function p_1 of T_1 is

$$p_1(t) = \lambda e^{-\lambda t}, \qquad t > 0, \tag{9.29}$$

as required. Since s was completely arbitrary, it could have coincided with the time of an event. It may be shown, in fact, that *the time interval between events is exponentially distributed with mean* $1/\lambda$. Since the average waiting time between events is $1/\lambda$, there are, roughly

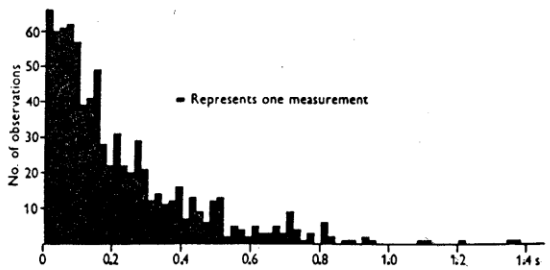

Figure 9.8. Histogram of time intervals between m.e.p.p.'s at the frog neuromuscular junction. [From Fatt and Katz (1952). Reproduced with the permission of The Physiological Society and the authors.]

speaking, on average λ events per unit time. Thus λ is called the mean rate or intensity.

Figure 9.8 shows the histogram, obtained by Fatt and Katz (1952), of time intervals between the spontaneous miniature endplate potentials at the frog neuromuscular junction. The histogram has the shape of an exponential distribution. Fatt and Katz were thus led to make their *Poisson hypothesis*, that the arrival times of the m.e.p.p.'s constituted a Poisson point process [see Van de Kloot, Kita, and Cohen (1975)].

The Poisson process as a primitive model for nerve-cell activity

Let the depolarization of a nerve cell be $\{V(t),\ t \geq 0\}$. Suppose that excitatory inputs occur at random in accordance with events in a simple Poisson process $\{N(t),\ t \geq 0\}$ with mean rate λ. Each excitatory input causes V to increase by a_E. When V reaches or exceeds the constant threshold level $\theta > 0$, the cell emits an action potential. Then

$$V(t) = a_E N(t), \qquad V < \theta,\ V(0) = 0. \tag{9.30}$$

In this primitive nerve-cell model, what is the probability distribution of the time interval between action potentials?

To answer this we first ask what is the waiting time T_k until the kth event in a simple Poisson process after the arbitrary time point s. We will show that *T_k has a gamma density with parameters k and λ.*

Proof. The kth event will occur in $(s+t, s+t+\Delta t]$ if and only if there are $k-1$ events in $(s, s+t]$ and one event in $(s+t, s+t+\Delta t]$.

It follows that

$$\Pr\{T_k \in (t, t+\Delta t]\} = \frac{e^{-\lambda t}(\lambda t)^{k-1}\lambda\,\Delta t}{(k-1)!} + o(\Delta t), \qquad k = 1, 2, \ldots. \tag{9.31}$$

Hence the density of T_k is

$$p_k(t) = \frac{\lambda(\lambda t)^{k-1}e^{-\lambda t}}{(k-1)!}, \qquad t > 0, \tag{9.32}$$

as required.

Hence we find that *the waiting time for the kth event has a gamma density with parameters k and λ.* Thus T_k has mean k/λ and variance k/λ^2. Some gamma densities are illustrated in Figure 9.9.

To return to the primitive nerve-cell model, an action potential is emitted when V reaches or exceeds θ, or, equivalently, when N reaches or exceeds θ/a_E. Letting $[x]$ denote the largest integer less than x we find that $1 + [\theta/a_E]$ excitatory inputs are required. *Hence the time interval between action potentials has a gamma density with parameters $1 + [\theta/a_E]$ and λ.* The mean time interval between action potentials is $(1 + [\theta/a_E])/\lambda$.

The gamma densities, in fact, resemble the ISI histograms obtained for many nerve cells and have often been fitted to them [see, for example, Stein (1965)]. However, the model employed to derive the gamma densities incorporates no decay of membrane potential between excitatory inputs. Only when the cell has an *exceedingly large*

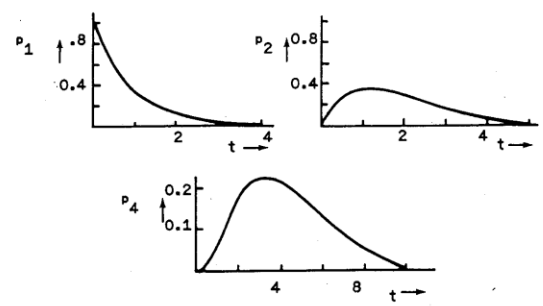

Figure 9.9. Gamma densities for $\lambda = 1$ and $k = 1, 2$, and 4. [Adapted from Cox and Miller (1965).]

time constant and the rate of incoming excitation is very fast will this approximation be valid. Finally, we note that when θ/a_E is large the gamma density becomes approximately that of a normal random variable.

9.5 Poisson excitation with Poisson inhibition

We will here consider another primitive model for nerve-cell activity in which excitation and inhibition arrive according to independent Poisson processes. This gives what is commonly called a *birth-and-death process*, which in this case Feller calls a *randomized random walk*. Much of what follows is taken from Feller's treatment (1966). Again we ignore the decay of membrane potential between inputs. Despite the unphysiological nature of the model, it is useful because:

(i) the model can be analyzed completely thereby providing a standard with which to compare other models and also real nerve cells; and

(ii) in a limiting case we obtain a Wiener process (see the next section), which is useful in many situations.

Let $V(t)$, $t \geq 0$, be the depolarization at time t. We assume that the number of excitatory inputs in $(0, t]$ is $N_E(t)$, where N_E is a Poisson process with mean rate λ_E, and that the number of inhibitory inputs is $N_I(t)$, where N_I is a Poisson process with mean rate λ_I. Each excitatory input makes V jump up by unity, whereas each inhibitory input causes V to jump down by unity. Thus

$$V(t) = N_E(t) - N_I(t), \qquad V(0) = 0, V < \theta. \tag{9.33}$$

The first thing we find is the probability distribution of $V(t)$ when there is no threshold for action potentials.

Consider what may happen in $(t, t+\Delta t]$. A jump in N_E occurs with probability $\lambda_E \Delta t + o(\Delta t)$, a jump in N_I occurs with probability $\lambda_I \Delta t + o(\Delta t)$, and with probabilities $1 - \lambda_E \Delta t + o(\Delta t)$ and $1 - \lambda_I \Delta t + o(\Delta t)$, respectively, N_E and N_I remain unchanged. The probability of a jump of either kind in V in $(t, t+\Delta t]$ is thus $(\lambda_E + \lambda_I)\Delta t + o(\Delta t)$ and the probability of no jump is $(1 - \lambda_E \Delta t + o(\Delta t))(1 - \lambda_I \Delta t + o(\Delta t)) = 1 - (\lambda_E + \lambda_I)\Delta t + o(\Delta t)$. The times at which V changes are thus a Poisson point process with mean rate $\lambda = \lambda_E + \lambda_I$. In fact, if we let $N(t)$ be the number of jumps (of either kind) of V in $(0, t]$, then

$$N(t) = N_E(t) + N_I(t), \tag{9.34}$$

and $\{N(t),\ t \geq 0\}$ is a Poisson process with mean rate λ. From the definition of conditional probability we find

$$\Pr(V \text{ jumps by } +1 \text{ in } (t, t+\Delta t] \mid \text{a jump in } V \text{ occurs in } (t, t+\Delta t]) = \lambda_E/\lambda \doteq p, \tag{9.35}$$

$$\Pr(V \text{ jumps by } -1 \text{ in } (t, t+\Delta t] \mid \text{a jump in } V \text{ occurs in } (t, t+\Delta]) = \lambda_I/\lambda \doteq q, \tag{9.36}$$

and $p + q = 1$.

We seek

$$p_m(t) = \Pr\{V(t) = m \mid V(0) = 0\}, \qquad m = 0, \pm 1, \pm 2, \ldots, \tag{9.37}$$

which is the conditional probability that $V(t) = m$ for an initial value zero. Such a quantity is an example of a *transition probability*, which we associate with a class of processes called *Markov* to which the process V belongs. We will for convenience drop the reference to the initial state and consider $m > 0$.

Let the process V be at m at t and suppose that $n \geq m$ jumps have occurred, of which n_1 were $+1$ and n_2 were -1. Then we must have

$$n = n_1 + n_2, \tag{9.38A}$$

$$m = n_1 - n_2, \tag{9.38B}$$

and hence

$$n_1 = (m+n)/2, \tag{9.38C}$$

$$n = m + 2n_2. \tag{9.38D}$$

The probability that $V(t) = m$ if n jumps have occurred in $(0, t]$ is the probability that a binomial random variable with parameters n and p takes the value n_1. That is,

$$\Pr\{V(t) = m \mid n \text{ jumps in } (0, t]\} = \binom{n}{n_1} p^{n_1} q^{n-n_1} = \binom{n}{\frac{n+m}{2}} p^{(n+m)/2} q^{(n-m)/2}. \tag{9.39}$$

By the law of total probability,

$$\Pr\{V(t) = m\} = \sum_{n \geq m} \Pr\{V(t) = m \mid n \text{ jumps in } (0, t]\} \times \Pr\{n \text{ jumps in } (0, t]\}. \tag{9.40}$$

Since n jumps in $(0, t]$ has probability $e^{-\lambda t}(\lambda t)^n/n!$, we find

$$p_m(t) = e^{-\lambda t} \sum_{n=m}^{\infty}{}' \frac{(\lambda t)^n}{n!} \binom{n}{\frac{n+m}{2}} p^{(n+m)/2} q^{(n-m)/2}, \quad (9.41)$$

where the prime on the summation sign indicates that summation is over either even or odd n depending on whether m is even or odd, respectively.

Utilizing (9.38) and the fact that $n = m, m+2, m+4, \ldots$ implies $n_2 = 0, 1, 2, \ldots$, this becomes

$$p_m(t) = e^{-\lambda t} \sum_{n_2=0}^{\infty} \frac{(\lambda t)^{m+2n_2}}{(m+2n_2)!} \binom{m+2n_2}{m+n_2} p^{m+n_2} q^{n_2}. \quad (9.42)$$

In terms of the modified Bessel function,

$$I_\rho(x) = \sum_{k=0}^{\infty} \frac{1}{k!\Gamma(k+\rho+1)} \left(\frac{x}{2}\right)^{2k+\rho}, \quad (9.43)$$

we get

$$p_m(t) = \left(\frac{\lambda_E}{\lambda_I}\right)^{m/2} e^{-\lambda t} I_m\left(2t\sqrt{\lambda_E \lambda_I}\right).$$

9.5.1 Time of first passage to threshold

We assume that there is a fixed threshold θ, which when reached by V leads to the emission of an action potential. The threshold condition is an imposed one and after a spike the potential is artificially reset to zero, possibly after a dead time or refractory period. Passage to time-varying thresholds in this model does not seem to have been considered.

We let θ be a positive integer and seek the *time of first passage of V to θ*, which is identified with the interspike interval. To find the probability distribution of the first-passage time, we employ the *method of images* in the symmetric case ($\lambda_E = \lambda_I$) and the *renewal equation* in the asymmetric case.

(A) *Symmetric case: method of images*

We will first find $p_m^*(t)$, the probability that the randomized random walk is at level m at t but has stayed below θ up to time t. This is in distinction to $p_m(t)$, which includes passages below, to, and above θ. Consider Figure 9.10, where a randomized walk process U is shown starting from the image point 2θ and having the value m at t. There is a one-to-one correspondence between such paths of U and

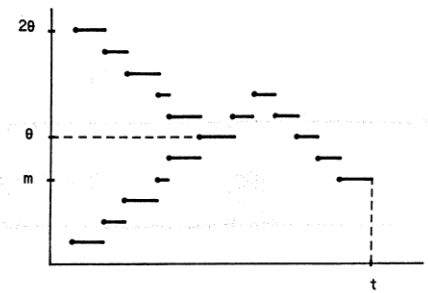

Figure 9.10. Paths considered in the method of images.

those of V that start at 0, touch and/or cross the level θ in (0, t), and end up at m at t. By symmetry the probability assigned to such paths is $p_{2\theta-m}(t)$. These paths are excluded in computing p_m^* so we have

$$p_m^*(t) = p_m(t) - p_{2\theta-m}(t). \quad (9.44)$$

To obtain the probability density $f_\theta(t)$ of the time of first passage to level θ, note that a first passage to θ occurs in $(t, t+\Delta t]$ if V has stayed below θ in $(0, t]$, is, in fact, at $\theta - 1$ at t, and a jump of $+1$ occurs in $(t, t+\Delta t]$. The probability of a jump of $+1$ in $(t, t+\Delta t]$ is $\lambda_E \Delta t = (\lambda/2)\Delta t$. Putting these probabilities together gives

$$f_\theta(t)\,\Delta t = p_{\theta-1}^*(t)(\lambda/2)\,\Delta t. \quad (9.45)$$

Utilizing (9.43) and (9.44) and the fact that $\lambda_E = \lambda_I$, we find

$$f_\theta(t) = \frac{\lambda}{2} e^{-\lambda t}\left[I_{\theta-1}(2\lambda_E t) - I_{\theta+1}(2\lambda_E t)\right], \qquad t > 0. \quad (9.46)$$

A more succinct expression results on using the recurrence relation for the modified Bessel function,

$$I_{\theta-1}(x) - I_{\theta+1}(x) = (2\theta/x) I_\theta(x), \quad (9.47)$$

whereupon

$$f_\theta(t) = (\theta/t) e^{-\lambda t} I_\theta(2\lambda_E t), \qquad t > 0. \quad (9.48)$$

(B) *General case: the renewal equation*

When the rates of arrival of jumps up and down are not equal, we resort to another method for obtaining $f_\theta(t)$. The idea on which the method is based is illustrated in Figure 9.11. A path is shown starting at zero and attaining the value $m > \theta$ at t. Since $m > \theta$,

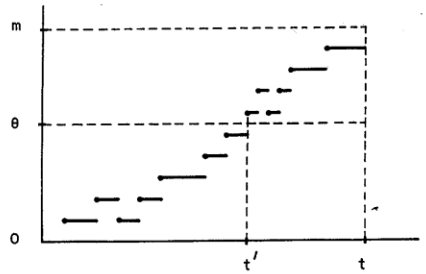

Figure 9.11. Paths that lead to the renewal equation.

such a path must have at some time before t passed through the level θ and, in particular, at some time $t' < t$ done this for the first time. Integrating over all such paths gives, using a continuous version of the law of total probability,

$$p_m(t) = \int_0^t f_\theta(t') p_{m-\theta}(t-t')\, dt'. \tag{9.49}$$

This integral equation is called a *renewal equation*, which we will solve using Laplace transform methods. Note that the probability of a transition from θ at t' to m at t is the same as the probability of a transition from 0 at time zero to $m-\theta$ at $t-t'$ (spatial and temporal homogeneity).

Equation (9.49) is a *convolution* of f_θ and $p_{m-\theta}$. From Table 3.2 we see that the Laplace transform of the convolution of two functions is the product of their Laplace transforms. Thus, denoting Laplace transforms by the extra subscript L,

$$p_{m,L}(s) = f_{\theta,L}(s) p_{m-\theta,L}(s), \qquad m < \theta, \tag{9.50}$$

where s is the transform variable. Rearranging this gives the following useful relation:

$$f_{\theta,L}(s) = p_{m,L}(s)/p_{m-\theta,L}(s). \tag{9.51}$$

The transforms on the right can be found as series. From (9.42) and Table 3.2, we find

$$\mathscr{L}\{e^{\lambda t} p_m(t)\} = \mathscr{L}\left\{\sum_{n_2=0}^{\infty} \frac{(\lambda t)^{m+2n_2}}{(m+2n_2)!} \binom{m+2n_2}{m+n_2} p^{m+n_2} q^{n_2}\right\}$$
$$= \frac{\lambda^m p^m}{s^{m+1}} \sum_{n_2=0}^{\infty} \left(\frac{\lambda pq}{s^2}\right)^{n_2} \binom{m+2n_2}{m+n_2}. \tag{9.52}$$

Utilizing the property $\mathscr{L}\{e^{ct} f(t)\} = f_L(s-c)$, we find

$$f_{\theta,L}(s) = \frac{\lambda^\theta p^\theta}{(s+\lambda)^\theta} \frac{\sum_{n_2=0}^{\infty} \left(\frac{\lambda pq}{(s+\lambda)^2}\right)^{n_2} \binom{m+2n_2}{m+n_2}}{\sum_{n_2=0}^{\infty} \left(\frac{\lambda pq}{(s+\lambda)^2}\right)^{n_2} \binom{m-\theta+2n_2}{m-\theta+n_2}}. \tag{9.53}$$

It is left as an exercise to show that this is the Laplace transform of the required first-passage time density,

$$f_\theta(t) = \theta \left(\frac{\lambda_E}{\lambda_I}\right)^{\theta/2} \frac{e^{-(\lambda_E+\lambda_I)t}}{t} I_\theta\left(2t\sqrt{\lambda_E \lambda_I}\right), \qquad t > 0. \tag{9.54}$$

Moments of the firing time

Let T_θ be the (random) time taken for V to reach θ from the initial state zero (resting state). Then T_θ has the probability density function f_θ. The nth moment of T_θ is

$$\mu_n = \int_0^\infty t^n f_\theta(t)\, dt. \tag{9.55}$$

When $n = 0$ we obtain the total probability mass of T_θ concentrated on $(0, \infty)$. That is,

$$\mu_0 = \Pr\{T_\theta < \infty\} = \int_0^\infty f_\theta(t)\, dt. \tag{9.56}$$

A series representation of this probability can be found but it is difficult to sum the series. However, by applying Theorem 7.1 of Karlin and Taylor (1975) on the probability of extinction in a general birth-and-death process, we find

$$\Pr\{T_\theta < \infty\} = \begin{cases} 1, & \lambda_E \geq \lambda_I, \\ \left(\frac{\lambda_E}{\lambda_I}\right)^\theta, & \lambda_E < \lambda_I \end{cases}. \tag{9.57}$$

Thus, if the mean rate of excitation is greater than or equal to the mean rate of inhibition, the time to reach threshold is finite with probability one. On the other hand, if the mean rate of inhibition is greater than that of excitation, the threshold may never be reached, which implies that the neuron may never fire an action potential. Note, however, that this result is obtained in a model that neglects the decay of potential between inputs.

If $\lambda_E < \lambda_I$ the mean firing time is infinite as T_θ has some probability mass at $t = \infty$. When $\lambda_E > \lambda_I$ the mean and variance of the firing

time can be found with the aid of the following relation (Gradshteyn and Ryzhik 1965, page 708):

$$\int_0^\infty e^{-\alpha x} I_\nu(\beta x)\,dx = \frac{\beta^\nu}{\sqrt{\alpha^2-\beta^2}\left(\alpha+\sqrt{\alpha^2-\beta^2}\right)^\nu}. \tag{9.58}$$

This yields

$$E[T_\theta] = \frac{\theta}{\lambda_E-\lambda_I}, \qquad \lambda_E > \lambda_I, \tag{9.59}$$

$$\mathrm{Var}[T_\theta] = \frac{\theta(\lambda_E+\lambda_I)}{(\lambda_E-\lambda_I)^3}, \ \lambda_E > \lambda_I. \tag{9.60}$$

The *coefficient of variation*, the standard deviation divided by the mean, is

$$\mathrm{CV}[T_\theta] = \left(\frac{\lambda_E+\lambda_I}{\theta(\lambda_E-\lambda_I)}\right)^{1/2}, \tag{9.61}$$

which indicates that the *coefficient of variation is inversely proportional to the square root of the threshold* for fixed rates of excitation and inhibition. When $\lambda_E = \lambda_I$, although $T_\theta < \infty$ with probability one, the mean (and higher-order moments) of T_θ is infinite.

Note that we have assumed that the excitatory and inhibitory jumps of V are of unit size. If instead the jumps have magnitude a, so that

$$V(t) = a[N_E(t) - N_I(t)], \tag{9.62}$$

then with a threshold $\theta > 0$, not necessarily an integer, the time to get to threshold will be the time for the process with unit jumps to reach $[1+\theta/a]$.

Tails of the firing-time density

Using the following asymptotic relation for the modified Bessel function at large arguments (Abramowitz and Stegun 1965, page 377),

$$I_\nu(x) \underset{x\to\infty}{\sim} \frac{e^x}{\sqrt{2\pi x}}\left\{1 - \frac{(4\nu^2-1)}{8x} + o\left(\frac{1}{x}\right)\right\}, \tag{9.63}$$

we deduce that when there is Poisson excitation and inhibition,

$$f_\theta(t) \underset{t\to\infty}{\sim} \frac{\theta}{2}\left(\frac{\lambda_E}{\lambda_I}\right)^{\theta/2} \frac{1}{\sqrt{\pi(\lambda_E\lambda_I)^{1/2}}} \frac{e^{-t(\sqrt{\lambda_E}-\sqrt{\lambda_I})^2}}{t^{3/2}}$$
$$\times\left\{1 - \frac{4\theta^2-1}{16t\sqrt{\lambda_E\lambda_I}} + o\left(\frac{1}{t}\right)\right\}, \tag{9.64}$$

whereas when there is Poisson excitation only, we have the exact result

$$f_\theta(t) = [\lambda^\theta/(\theta-1)!]\,t^{\theta-1}e^{-\lambda t}. \tag{9.65}$$

Thus the density of the first-passage time to level θ is quite different in its tails, depending on the presence or absence of inhibition.

9.6 The Wiener process

We will soon proceed to more realistic models, which incorporate the decay of membrane potential between synaptic inputs. Before doing so, we consider an approximation to the process V defined in the previous section. The approximating process is a *Wiener process* (*or Brownian motion*), which belongs to the general class of Markov processes called *diffusion processes*. These general concepts will be explained later. Gerstein and Mandelbrot (1964) pioneered the use of the Wiener process in neural modeling.

Diffusion processes have trajectories that are continuous functions of t, in distinction to the randomized random walk whose sample paths are discontinuous. The study of diffusion processes is often less difficult than that of their discontinuous counterparts chiefly because the equations describing their important properties are differential equations about which more is known than differential-difference equations, which arise for discontinuous processes. Among the reasons for studying the Wiener process as a model for nerve membrane potential are:

(i) it is a thoroughly studied process and many of the relevant mathematical problems have been solved; and
(ii) from the Wiener process we may construct many other more realistic models of nerve-cell activity.

The Wiener process as a limiting case of a random walk

Consider the process defined by

$$V_a(t) = a[N_E(t) - N_I(t)], \qquad t \ge 0, \tag{9.66}$$

where a is a constant, and N_E and N_I are independent Poisson processes with mean rates $\lambda_E = \lambda_I = \lambda$. The process V_a has jumps up and down of magnitude a. We note that

$$E[V_a(t)] = a[\lambda_E - \lambda_I]t = 0, \tag{9.67}$$

$$\mathrm{Var}[V_a(t)] = a^2[\mathrm{Var}[N_E(t)] + \mathrm{Var}[N_I(t)]] = 2a^2\lambda t. \tag{9.68}$$

The characteristic function of $V_a(t)$ is

$$\begin{aligned}\phi_a(u;t) &= E[\exp(iuV_a(t))] \\ &= E[\exp(iua(N_E(t)-N_I(t)))] \\ &= E[\exp(iuaN_E(t))]\,E[\exp(-iuaN_I(t))] \qquad (9.69)\end{aligned}$$

by the independence of N_E and N_I. From Section 9.2 we find

$$\phi_a(u;t) = \exp\{\lambda t(e^{iua}+e^{-iua}-2)\}. \qquad (9.70)$$

To standardize the random variables $V_a(t)$, we let

$$\lambda = 1/2a^2, \qquad (9.71)$$

so that $V_a(t)$ has mean zero and variance t for all a. It is left as an exercise to show that

$$\lim_{a\to 0}\phi_a(u;t) = e^{-\frac{1}{2}u^2t} \doteq \phi(u;t). \qquad (9.72)$$

Thus, from Section 9.2, we see that $\phi(u;t)$ is the characteristic function of a normal random variable with mean zero and variance t.

One way to characterize the distance between two random variables is by the differences between their distribution functions. Let $\{X_n, n=1,2,\dots\}$ be a sequence of random variables with distribution function F. If

$$\lim_{n\to\infty} F_n(x) = F(x),$$

for all points x at which F is continuous, we say the sequence $\{X_n\}$ *converges in distribution* to X. We write

$$X_n \xrightarrow{d} X.$$

A basic theorem of probability theory [see, for example, Ash (1970), page 171] tells us that to establish convergence in distribution it is sufficient to prove convergence of the corresponding sequence of characteristic functions. Hence as $a\to 0$, the sequence of random variables $V_a(t)$ converges in distribution to a normal random variable with mean zero and variance t. We let this limiting variable be $W(t)$,

$$V_a(t) \xrightarrow[a\to 0]{} W(t). \qquad (9.73)$$

The process $\{W(t), t\ge 0\}$ is called a *standard Wiener process* and more will be said about the convergence of V_a to W in Section 9.9.

The "standard" refers to the values of the mean and variance,

$$E[W(t)] = 0, \qquad (9.74)$$

$$\mathrm{Var}[W(t)] = t. \qquad (9.75)$$

Definition and some properties of W

In the above limiting procedure, we obtained W by letting the jump size in the random walk become smaller, whereas the rates at which the jumps arrived became faster. Thus W has sample paths, which are, in fact, continuous. So far we have only considered $W(t)$, the value of W at time t. A definition of the process $\{W(t), t\ge 0\}$ is the following.

Definition

$\{W(t), t\ge 0\}$ is a standard Wiener process (Brownian motion) if:

(a) $W(0)=0$;
(b) given any $0\le t_0<t_1<t_2<\cdots<t_{n-1}<t_n$, the random variables $W(t_k)-W(t_{k-1})$, $k=1,2,\dots,n$, are independent; and
(c) for any $0\le t_1<t_2$, $W(t_2)-W(t_1)$ is a normal random variable with mean zero and variance t_2-t_1.

Note that $\{V_a(t)\}$ satisfies (a), (b), and (c) asymptotically as $a\to 0$.

Thus W shares with the Poisson process the property of having stationary independent increments. The density of $W(t)$ is

$$f_W(x,t) = \frac{1}{\sqrt{2\pi t}}\exp\left\{-\frac{x^2}{2t}\right\}, \qquad -\infty<x<\infty,\ t>0. \qquad (9.76)$$

An attempt is made in Figure 9.12 to depict some sample paths for W. Although the paths of W are smooth enough to be continuous, they

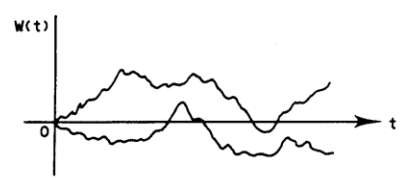

Figure 9.12. A sketch of two sample paths of W.

are, with probability one, nondifferentiable. Nevertheless, the "derivative" of W, denoted by w, is called *white noise* and is a useful concept. Whenever w appears in an equation, an integration is implied.

Wiener process with drift

We may construct new processes from W by multiplying it by a constant and adding a linear drift. Thus

$$X(t) = x_0 + \sigma W(t) + \mu t, \qquad t > 0, \tag{9.77}$$

where $X(0) = x_0$ defines a *Wiener process with variance parameter* σ *and drift parameter* μ. Linear operations on Gaussian (normal) processes, such as W, produce Gaussian processes. Since

$$E[X(t)] = x_0 + \mu t, \tag{9.78}$$

$$\mathrm{Var}[X(t)] = \sigma^2 t, \tag{9.79}$$

the density of $X(t)$ is

$$f_X(x,t) = \frac{1}{\sqrt{2\pi\sigma^2 t}} \exp\left\{-\frac{(x - x_0 - \mu t)^2}{2\sigma^2 t}\right\}, \qquad -\infty < x < \infty,\ t > 0. \tag{9.80}$$

As $t \to 0^+$, $f_X(x,t)$ approaches $\delta(x - x_0)$.

The nerve-cell model

It is the Wiener process with drift that Gerstein and Mandelbrot (1964) employed as an approximate model for nerve membrane potential. Roughly speaking, the following correspondences prevail between the original random-walk and its smoothed version. If excitation and inhibition arrive at the occurrence times of jumps in the two independent Poisson processes N_E and N_I, with mean rates λ_E and λ_I, and each excitatory input causes the depolarization $V(t)$ to jump up by $a_E \geq 0$ whereas each inhibitory input causes $V(t)$ to jump down by $a_I \geq 0$ units, then

$$V(t) = a_E N_E(t) - a_I N_I(t), \qquad t \geq 0. \tag{9.81}$$

To obtain an approximation to V with continuous sample paths, we use a Wiener process with drift that has the same mean and variance as V. Thus in (9.77) we put

$$\mu = a_E \lambda_E - a_I \lambda_I, \tag{9.82}$$

$$\sigma = \sqrt{a_E^2 \lambda_E + a_I^2 \lambda_I}. \tag{9.83}$$

We have, of course, left something out as we have not performed any limiting operations on the jump amplitudes or the mean rates to obtain the continuous process from the discontinuous one. We have made what has been called the usual diffusion approximation (Walsh 1981a) in the hope that the original process and the approximating smooth process go to the same places at about the same times.

9.6.1 First-passage time to threshold

We again endow our model with a threshold condition, namely, that an action potential occurs when the approximating process X reaches the level θ, assumed constant, for the first time. The firing time is then the random variable

$$T_\theta = \inf\{t \mid X(t) = \theta\}, \qquad X(0) = x_0 < \theta. \tag{9.84}$$

The density of T_θ can be found in closed form, first by the method of images when there is no drift, and second by the renewal-equation approach when drift is present.

(A) *The drift* $\mu = 0$: *method of images*

We use essentially the same argument as for the symmetric randomized random walk. The *transition probability density* function of X is defined as

$$p(x, t \mid x_0) = d/dx \Pr\{X(t) \leq x \mid X(0) = x_0\}, \tag{9.85}$$

and for the unrestricted process this is given by (9.80). We have, roughly speaking,

$$\Pr\{X(t) \in (x, x + dx] \mid X(0) = x_0\} \simeq p(x, t \mid x_0)\, dx. \tag{9.86}$$

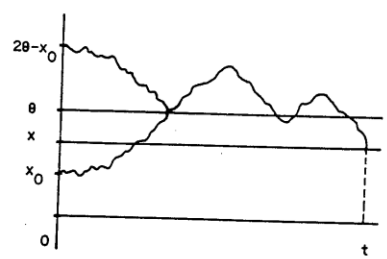

Figure 9.13. Paths of X for the method of images

Referring to Figure 9.13, consider a path starting at x_0 and touching or rising above the level θ in $(0, t)$. For each such path there is one from the image point $2\theta - x_0$ to x. Hence

$$\Pr\{X(t) \in (x, x+dx] \text{ and } X(t') < \theta \text{ for all } t' \in (0,t) \mid X(0) = x_0\}.$$
$$= [p(x,t|x_0) - p(x,t|2\theta - x_0)]\,dx. \tag{9.87}$$

Integrating over all x less than θ and noting that if X has stayed below θ up to t, then $T_\theta > t$,

$$\Pr\{T_\theta > t\} = \Pr\left\{\max_{0 \le t' \le t} X(t') < \theta \mid X(0) = x_0\right\}$$
$$= \int_{-\infty}^{\theta} [p(x,t|x_0) - p(x,t|2\theta - x_0)]\,dx. \tag{9.88}$$

Now (again roughly speaking) transitions from $2\theta - x_0$ to x occur with the same probabilities as transitions from x_0 to $2\theta - x$, so $p(x,t|2\theta - x_0) = p(2\theta - x, t|x_0)$. Thus, since $\Pr\{T_\theta > t\} = 1 - \Pr\{T_\theta \le t\}$, $t > 0$,

$$\Pr\{T_\theta \le t\} = 1 - \int_{-\infty}^{\theta} [p(x,t|x_0) - p(2\theta - x, t|x_0)]\,dx. \tag{9.89}$$

Changing variables in the second integral yields

$$\Pr\{T_\theta \le t\} = 2\int_{\theta}^{\infty} p(x,t|x_0)\,dx$$
$$= \sqrt{\frac{2}{\pi\sigma^2 t}} \int_{\theta}^{\infty} \exp\left\{-\frac{(x-x_0)^2}{2\sigma^2 t}\right\} dx. \tag{9.90}$$

With the additional change of variable $z = (x - x_0)/\sigma\sqrt{t}$, we get

$$\Pr\{T_\theta \le t\} = \sqrt{\frac{2}{\pi}} \int_{(\theta - x_0)/\sigma\sqrt{t}}^{\infty} e^{-z^2/2}\,dz. \tag{9.91}$$

But the distribution function of T_θ is $F_\theta(t) = \Pr\{T_\theta \le t\}$, and its derivative is the density f_θ of T_θ. Differentiating (9.91), we obtain the following expression for the *density of the first-passage time to level θ for the driftless Wiener Process*:

$$f_\theta(t) = \frac{\theta - x_0}{\sqrt{2\pi\sigma^2 t^3}} \exp\left\{-\frac{(\theta - x_0)^2}{2\sigma^2 t}\right\}, \qquad t > 0,\ \theta > x_0. \tag{9.92}$$

Letting $t \to \infty$ in (9.91), we see immediately that passage to level θ is certain.

(B) *The general case including $\mu \neq 0$: renewal equation*

Using the same argument as for the general randomized random walk, we obtain the renewal equation

$$p(x,t|x_0) = \int_0^t f_\theta(t')\,p(x, t - t'|\theta)\,dt', \qquad x > \theta. \tag{9.93}$$

Taking Laplace transforms and rearranging gives

$$f_{\theta,L}(s) = \frac{p_L(x,s|x_0)}{p_L(x,s|\theta)}. \tag{9.94}$$

With the aid of the standard transform (Abramowitz and Stegun 1965, page 1026),

$$\mathscr{L}\left\{\frac{1}{\sqrt{\pi t}} \exp\left(-\frac{k^2}{4t}\right)\right\} = \frac{1}{\sqrt{s}} e^{-k\sqrt{s}}, \qquad k \ge 0, \tag{9.95}$$

we find

$$p_L(x,s|x_0) = \frac{1}{\sqrt{2\sigma^2}} \exp\left[\frac{\mu(x-x_0)}{\sigma^2}\right] \frac{\exp\left[-\dfrac{(x-x_0)}{\sigma}\sqrt{2(s-c)}\right]}{\sqrt{s-c}}, \tag{9.96}$$

where $c = -\mu^2/2\sigma^2$. It is left as an exercise to show that the *Laplace transform of the first-passage time density* is

$$f_{\theta,L}(s) = \exp\left\{\frac{(\theta - x_0)}{\sigma^2}\left(\mu - \sqrt{\mu^2 + 2\sigma^2 s}\right)\right\}. \tag{9.97}$$

The inversion of this transform is facilitated by another standard result (Abramowitz and Stegun 1965, page 1026),

$$\mathscr{L}\left\{\frac{k}{2\sqrt{\pi t^3}} \exp\left(-\frac{k^2}{4t}\right)\right\} = \exp(-k\sqrt{s}), \qquad k \ge 0, \tag{9.98}$$

which yields the *inverse Gaussian* density:

$$f_\theta(t) = \frac{\theta - x_0}{\sqrt{2\pi\sigma^2 t^3}} \exp\left[-\frac{(\theta - x_0 - \mu t)^2}{2\sigma^2 t}\right], \qquad t > 0,\ \theta > x_0. \tag{9.99}$$

Moments of the firing time

We may find the probability that X ever reaches θ by utilizing the relation

$$f_{\theta,L}(0) = \int_0^\infty f_\theta(t)\,dt = \Pr\{T_\theta < \infty\}. \tag{9.100}$$

From (9.97)

$$f_{\theta,L}(0) = \exp\left[\frac{(\theta - x_0)}{\sigma^2}(\mu - |\mu|)\right], \tag{9.101}$$

since $\sqrt{\mu^2} = |\mu|$ must be nonnegative. Since $|\mu| = \mu$ if $\mu \geq 0$ and $|\mu| = -\mu$ if $\mu < 0$, we have

$$\Pr\{T_\theta < \infty\} = \begin{cases} 1, & \mu \geq 0, \\ \exp\left[-\dfrac{2|\mu|(\theta - x_0)}{\sigma^2}\right], & \mu < 0. \end{cases} \tag{9.102}$$

Thus, if the drift is zero or toward the threshold, an action potential is generated in a finite time with probability one. On the other hand, as with the random-walk model, if the drift is away from the barrier so that $a_E\lambda_E < a_I\lambda_I$, there is probability $1 - \exp[-2|\mu|(\theta - x_0)/\sigma^2]$ that no action potential is ever generated (an eternally silent cell).

In the case $\mu \geq 0$ the mean waiting time for the occurrence of an action potential can be found from

$$E[T_\theta] = -\left.\frac{df_{\theta,L}(s)}{ds}\right|_{s=0}, \tag{9.103}$$

and the second moment can be found from

$$E[T_\theta^2] = \left.\frac{d^2 f_{\theta,L}(s)}{ds^2}\right|_{s=0}. \tag{9.104}$$

It is an exercise to show that

$$E[T_\theta] = \frac{\theta - x_0}{\mu}, \tag{9.105}$$

$$\mathrm{Var}[T_\theta] = \frac{(\theta - x_0)\sigma^2}{\mu^3}, \qquad \mu > 0,\ \theta \geq x_0. \tag{9.106}$$

When $\mu = 0$ the first- and higher-order moments of T_θ are infinite, as they must also be when $\mu < 0$.

In terms of the original physiological parameters of the model, assuming the initial value of the membrane potential is resting level,

the mean ISI is

$$E[T_\theta] = \frac{\theta}{a_E\lambda_E - a_I\lambda_I}, \tag{9.107}$$

and the variance is

$$\mathrm{Var}[T_\theta] = \frac{\theta(a_E^2\lambda_E + a_I^2\lambda_I)}{(a_E\lambda_E - a_I\lambda_I)^3}, \qquad a_E\lambda_E > a_I\lambda_I. \tag{9.108}$$

This gives a coefficient of variation of

$$\mathrm{CV}[T_\theta] = \left(\frac{a_E^2\lambda_E + a_I^2\lambda_I}{2\theta(a_E\lambda_E - a_I\lambda_I)^3}\right)^{1/2}. \tag{9.109}$$

Again, for fixed values of the remaining parameters, the coefficient of variation of the ISI is inversely proportional to the square root of the threshold.

A numerical example

With time in units of the membrane time constant and voltages in millivolts, we will find the mean, variance, coefficient of variation, and density of the firing time for the following parameter values: $\theta = 10\,\mathrm{mV}$, $a_E = a_I = 1\,\mathrm{mV}$, $\lambda_E = 2.5$, and $\lambda_I = 0.5$. Then

$$E[T_\theta] = 5,$$

$$\mathrm{Var}[T_\theta] = \tfrac{15}{8},$$

$$\mathrm{CV}[T_\theta] = 0.27.$$

The density of T_θ is sketched in Figure 9.14.

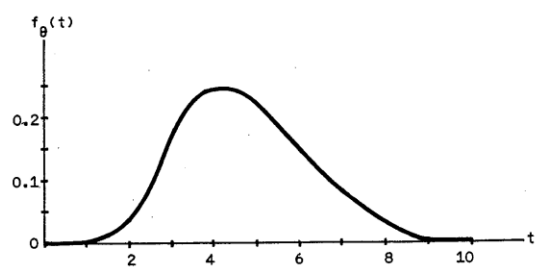

Figure 9.14. Density of the firing time for the Wiener process with drift model with parameters as given in the text.

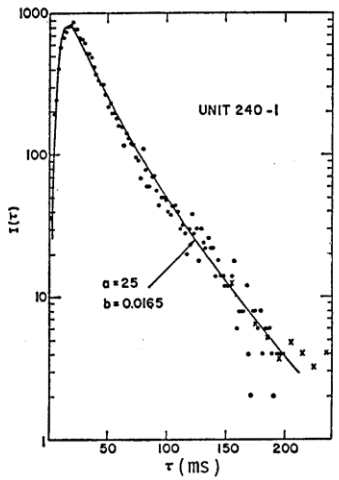

Figure 9.15. The fitting of an experimental ISI histogram to the first-passage time density of a Wiener process with drift. [From Gerstein and Mandelbrot (1964). Reproduced from *The Biophysical Journal* by copyright permission of The Biophysical Society.]

Although the Wiener process with drift has a poor physiological foundation as a model for nerve-cell membrane potential, formula (9.99) for the first-passage time density has been successfully fitted to the ISI histograms of real neurons. For example, Gerstein and Mandelbrot (1964) put $f_\theta(t)$ in the form

$$f_\theta(t) = Kt^{-3/2}\exp(-a/t - bt), \tag{9.110}$$

regarding a and b as parameters with K determined by the normalization condition $\int_0^\infty f_\theta(t)\,dt = 1$. One case of their fitting procedure is shown in Figure 9.15 for a cell in the cat cochlear nucleus. Although the agreement between experimental and theoretical firing-time distributions is excellent in this and other cases, the fitting procedure is of limited use. The model itself does not incorporate the realities of nerve-cell behavior, and the parameters are not related to the physiological variables.

9.7 Markov processes

In the following few sections we will need some basic results from the theory of *Markov processes*. This section contains a brief review of the pertinent results. Technicalities are omitted as only the applications are needed. For advanced mathematical details and foundations see such books as those of Dynkin (1965), Breiman (1968), and Gihman and Skorohod (1972, 1974, 1975, 1976).

To say that $\{X(t); t \geq 0\} \doteq X$ is a Markov process is to say that if X is known at time s, then the probabilities that X takes on its various possible values at any future time are completely determined.

Transition probability functions

Markov processes are characterized by their *transition probability functions*. Let $\{X(t), t \geq 0\}$ be a Markov process in continuous time. Suppose at time s, X is known to have the value x. We call

$$P(y, t|x, s) = \Pr\{X(t) \leq y | X(s) = x\}, \qquad s \leq t, \tag{9.111}$$

the *transition probability distribution function*.

If P is differentiable in y, its derivative is the *transition probability density function*

$$p(y, t|x, s) = \partial P/\partial y(y, t|x, s). \tag{9.112}$$

Processes for which the transition probabilities depend only on *time differences* $t - s$ are called *temporally homogeneous*. All of the random processes considered thus far in this chapter, being processes with independent increments, are Markov processes and are also temporally homogeneous. We will only be concerned in the sequel with temporally homogeneous Markov processes.

Example

The Wiener process with drift parameter μ and variance parameter σ^2 considered in the previous section has the transition probability distribution function

$$P(y, t|x, s) = \frac{1}{\sqrt{2\pi\sigma^2(t-s)}} \int_{-\infty}^{y} \exp\left[-\frac{(z - x - \mu(t-s))}{2\sigma^2(t-s)}\right] dz, \tag{9.113}$$

and transition probability density function

$$p(y, t|x, s) = \frac{1}{\sqrt{2\pi\sigma^2(t-s)}} \exp\left[-\frac{(y - x - \mu(t-s))^2}{2\sigma^2(t-s)}\right]. \tag{9.114}$$

Infinitesimal generators

Markov processes admit a characterization through operators that describe the changes in the process in small time intervals. Let X be a Markov process in continuous time. Then the *infinitesimal operator*, or *infinitesimal generator*, of X is defined for suitable functions f through

$$(\mathscr{A}f)(x) = \lim_{\Delta t \downarrow 0} \frac{E[f(X(t+\Delta t)) - f(X(t)) | X(t) = x]}{\Delta t} \tag{9.115}$$

The infinitesimal operator of a process can be calculated from a knowledge of its transition probability function. In the theory of Markov processes it is shown how the reverse step can be carried out through the Kolmogorov partial differential equations.

Examples

(i) *Poisson process.* For the Poisson process, if $X(t) = x$, then $X(t + \Delta t) = x + 1$ with probability $\lambda \Delta t + o(\Delta t)$ and $X(t + \Delta t) = x$ with probability $1 - \lambda \Delta t + o(\Delta t)$. Thus

$$\begin{aligned} &E[f(X(t+\Delta t)) - f(X(t)) | X(t) = x] \\ &\quad = \lambda \Delta t f(x+1) + (1 - \lambda \Delta t) f(x) - f(x) + o(\Delta t). \end{aligned}$$

Putting this in (9.114) and carrying out the limiting operation gives

$$(\mathscr{A}f)(x) = \lambda[f(x+1) - f(x)].$$

(ii) *Randomized random walk.* A similar calculation shows that for the process of Equation (9.81)

$$(\mathscr{A}f)(x) = \lambda_E f(x + a_E) + \lambda_I f(x - a_I) - (\lambda_E + \lambda_I) f(x).$$

(iii) *Wiener process with drift.* For the Wiener process with drift, use of (9.114) leads to

$$(\mathscr{A}f)(x) = \frac{\sigma^2}{2} \frac{d^2 f}{dx^2} + \mu \frac{df}{dx}.$$

Diffusion processes

Roughly speaking, diffusion processes are continuous-time Markov processes whose sample paths are continuous. Such processes are characterized by their infinitesimal mean

$$\alpha(x) = \lim_{\Delta t \downarrow 0} \frac{E[X(t+\Delta t) - X(t) | X(t) = x]}{\Delta t}, \tag{9.116}$$

and infinitesimal variance

$$\beta^2(x) = \lim_{\Delta t \downarrow 0} \frac{\mathrm{Var}[X(t+\Delta t) - X(t) | X(t) = x]}{\Delta t} \tag{9.117}$$

The infinitesimal generator of such a process is

$$(\mathscr{A}f)(x) = \alpha(x) \frac{df}{dx} + \frac{\beta^2(x)}{2} \frac{d^2 f}{dx^2}. \tag{9.118}$$

Diffusion processes can be described by their *stochastic differential equations*. The process with infinitesimal generator (9.118) has the stochastic differential

$$dX = \alpha(X)\, dt + \beta(X)\, dW, \tag{9.119}$$

where W is a standard Wiener process. The differential relation is defined by its integral version

$$X(t) = X_0 + \int_0^t \alpha(X(t'))\, dt' + \int_0^t \beta(X(t'))\, dW(t'), \tag{9.120}$$

where the initial condition is $X(0) = X_0$.

The first integral in (9.120) is a Riemann integral but the second is a *stochastic integral* with respect to W (Itô 1951; Stratonovich 1966). This integral is defined and some of its properties are given, for example, in Jaswinski (1970).

Processes with jumps

Consider a Markov random process Y that has jumps of various magnitudes. Let $\nu(t, A)$ record the number of jumps of Y up to time t that have magnitudes in the set A. Suppose for fixed A, $\nu(t, A)$ is a temporally homogeneous Poisson process with mean rate $\Pi(A)$ depending on A. Then $E[\nu(t, A)] = t\Pi(A)$. Suppose further that if A_i, $i = 1, 2, \ldots, n$, are disjoint sets, then $\nu(t, A_1)$, $\nu(t, A_2), \ldots, \nu(t, A_n)$ are mutually independent.

If we integrate over all possible jump amplitudes, we recover the original process Y,

$$Y(t) = \int_{\mathbf{R}} u \nu(t, du), \tag{9.121}$$

and the total jump rate (i.e., mean rate of jumps of all magnitudes) is

$$\Lambda = \int_{\mathbf{R}} \Pi(du). \tag{9.122}$$

The process Y is a *compound Poisson process*. If the *rate measure* Π

has a density so that $\Pi(du) = \pi(u)\,du$, then

$$\Lambda = \int_{\mathbf{R}} \pi(u)\,du. \tag{9.123}$$

Example

For the randomized random walk, jumps of magnitude $+a_E$ and $-a_I$ occur with mean rates λ_E and λ_I, respectively. Then

$$\pi(u) = \lambda_E \delta(u - a_E) + \lambda_I \delta(u + a_I)$$

and the total mean rate of jumps is

$$\Lambda = \int_{\mathbf{R}} \pi(u)\,du = \lambda_E + \lambda_I.$$

A stochastic differential equation can be written down, which describes a general Markov process with diffusion and jump components

$$dX = \alpha(X)\,dt + \beta(X)\,dW + \int_{\mathbf{R}} \gamma(X, u)\nu(dt, du). \tag{9.124}$$

Again the differential is an abbreviation for the integral equation

$$X(t) = X(0) + \int_0^t \alpha(X(t'))\,dt' + \int_0^t \beta(X(t'))\,dW(t')$$
$$+ \int_0^t \int_{\mathbf{R}} \gamma(X(t'), u)\nu(dt', du), \tag{9.125}$$

where the third integral is a stochastic integral with respect to ν. It may be shown (Gihman and Skorohod 1972) that the infinitesimal generator of the process defined by (9.125) is

$$(\mathscr{A}f)(x) = \alpha(x)\frac{df}{dx} + \frac{\beta^2(x)}{2}\frac{d^2f}{dx^2} + \int_{\mathbf{R}} f(x + \gamma(x, u))\Pi(du) - \Lambda f. \tag{9.126}$$

First-exit times

The part of the theory that concerns us most is the theory of exit times as these directly relate to the random firing of neurons. Suppose $X(0) = x$ where $a < x < b$. The *first-exit time* of X from the interval (a, b) is the random variable

$$T_{ab}(x) = \inf\{t \mid X(t) \notin (a, b)\}, \qquad X(0) = x \in (a, b). \tag{9.127}$$

Let the distribution function of $T_{ab}(x)$ be

$$F_{ab}(x, t) = \Pr\{T_{ab}(x) \le t\}. \tag{9.128}$$

Then (Tuckwell 1976a) F_{ab} can be found as the solution of

$$\frac{\partial F_{ab}}{\partial t} = \mathscr{A}F_{ab}, \qquad x \in (a, b),\ t > 0, \tag{9.129}$$

where $\mathscr{A}$ is a partial differential–integral operator, the infinitesimal operator of x. The initial condition is

$$F_{ab}(x, 0) = \begin{cases} 0, & x \in (a, b), \\ 1, & x \notin (a, b), \end{cases} \tag{9.130}$$

and with boundary conditions

$$F_{ab}(x, t) = 1, \qquad x \notin (a, b),\ t \ge 0. \tag{9.131}$$

Differentiating F_{ab} with respect to t, we get the density of the first-exit time $f_{ab}(x, t)$, which satisfies the same equation (9.129) as F_{ab} with boundary conditions

$$f_{ab}(x, t) = \delta(t), \qquad x \notin (a, b), \tag{9.132}$$

$$f_{ab}(x, 0) = 0, \qquad x \in (a, b). \tag{9.133}$$

In the case of diffusion processes, where exit from (a, b) is attained by hitting either a or b, the condition $x \notin (a, b)$ can be replaced by $x = a$ or $x = b$.

The *moments* of the first-exit time are defined through

$$\mu_{n,ab}(x) = \int_0^\infty t^n f_{ab}(x, t)\,dt, \qquad n = 0, 1, 2, \ldots. \tag{9.134}$$

These satisfy the recursion system of equations

$$\mathscr{A}\mu_{n,ab} = -n\mu_{n-1,ab}, \qquad x \in (a, b). \tag{9.135}$$

When $n = 0$, one obtains the probability $\mu_{0,ab}$ that X ever leaves (a, b) and for this the boundary condition is

$$\mu_{0,ab}(x) = 1, \qquad x \notin (a, b). \tag{9.136}$$

For the first- and higher-order moments of $T_{ab}(x)$, the boundary conditions are

$$\mu_{n,ab}(x) = 0, \qquad x \notin (a, b),\ n = 1, 2, \ldots, \tag{9.137}$$

and $\mu_{n,ab}$ is bounded on (a, b), if (a, b) is a finite interval. Exit times from intervals such as $(-\infty, b)$ are obtained by letting $a \to -\infty$ in the results for finite intervals.