1 Neural Encoding I: Firing Rates and Spike Statistics

1.1   Introduction

Neurons are remarkable among the cells of the body in their ability to propagate signals rapidly over large distances. They do this by generating characteristic electrical pulses called action potentials or, more simply, spikes that can travel down nerve fibers. Neurons represent and transmit information by firing sequences of spikes in various temporal patterns. The study of neural coding, which is the subject of the first four chapters of this book, involves measuring and characterizing how stimulus attributes, such as light or sound intensity, or motor actions, such as the direction of an arm movement, are represented by action potentials.

The link between stimulus and response can be studied from two opposite points of view. Neural encoding, the subject of chapters 1 and 2, refers to the map from stimulus to response. For example, we can catalog how neurons respond to a wide variety of stimuli, and then construct models that attempt to predict responses to other stimuli. Neural decoding refers to the reverse map, from response to stimulus, and the challenge is to reconstruct a stimulus, or certain aspects of that stimulus, from the spike sequences it evokes. Neural decoding is discussed in chapter 3. In chapter 4, we consider how the amount of information encoded by sequences of action potentials can be quantified and maximized. Before embarking on this tour of neural coding, we briefly review how neurons generate their responses and discuss how neural activity is recorded. The biophysical mechanisms underlying neural responses and action potential generation are treated in greater detail in chapters 5 and 6.

Properties of Neurons

Neurons are highly specialized for generating electrical signals in response to chemical and other inputs, and transmitting them to other cells. Some important morphological specializations, seen in figure 1.1, are the dendrites that receive inputs from other neurons and the axon that carries the neuronal output to other cells. The elaborate branching structure of

axons and dendrites

the dendritic tree allows a neuron to receive inputs from many other neurons through synaptic connections. The cortical pyramidal neuron of figure 1.1A and the cortical interneuron of figure 1.1C each receive thousands of synaptic inputs, and for the cerebellar Purkinje cell of figure 1.1B the number is over 100,000. Figure 1.1 does not show the full extent of the axons of these neurons. Axons from single neurons can traverse large fractions of the brain or, in some cases, of the entire body. In the mouse brain, it has been estimated that cortical neurons typically send out a total of about 40 mm of axon and have approximately 10 mm of total dendritic cable in their branched dendritic trees. The axon makes an average of 180 synaptic connections with other neurons per mm of length and the dendritic tree receives, on average, 2 synaptic inputs per μm. The cell body or soma of a typical cortical neuron ranges in diameter from about 10 to 50 μm.

ion channels

Along with these morphological features, neurons have physiological specializations. Most prominent among these are a wide variety of membrane-spanning ion channels that allow ions, predominantly sodium (Na^+), potassium (K^+), calcium (Ca^{2+}), and chloride (Cl^-), to move into and out of the cell. Ion channels control the flow of ions across the cell membrane by opening and closing in response to voltage changes and to both internal and external signals.

membrane potential

hyperpolarization and depolarization

The electrical signal of relevance to the nervous system is the difference in electrical potential between the interior of a neuron and the surrounding extracellular medium. Under resting conditions, the potential inside the cell membrane of a neuron is about -70 mV relative to that of the surrounding bath (which is conventionally defined to be 0 mV), and the cell is said to be polarized. Ion pumps located in the cell membrane maintain concentration gradients that support this membrane potential difference. For example, Na^+ is much more concentrated outside a neuron than inside it, and the concentration of K^+ is significantly higher inside the neuron than in the extracellular medium. Ions thus flow into and out of a cell due to both voltage and concentration gradients. Current in the form of positively charged ions flowing out of the cell (or negatively charged ions flowing into the cell) through open channels makes the membrane potential more negative, a process called hyperpolarization. Current flowing into the cell changes the membrane potential to less negative or even positive values. This is called depolarization.

action potential

refractory period

If a neuron is depolarized sufficiently to raise the membrane potential above a threshold level, a positive feedback process is initiated, and the neuron generates an action potential. An action potential is a roughly 100 mV fluctuation in the electrical potential across the cell membrane that lasts for about 1 ms (figure 1.2A). Action potential generation also depends on the recent history of cell firing. For a few milliseconds just after an action potential has been fired, it may be virtually impossible to initiate another spike. This is called the absolute refractory period. For a longer interval known as the relative refractory period, lasting up to tens of milliseconds after a spike, it is more difficult to evoke an action potential.
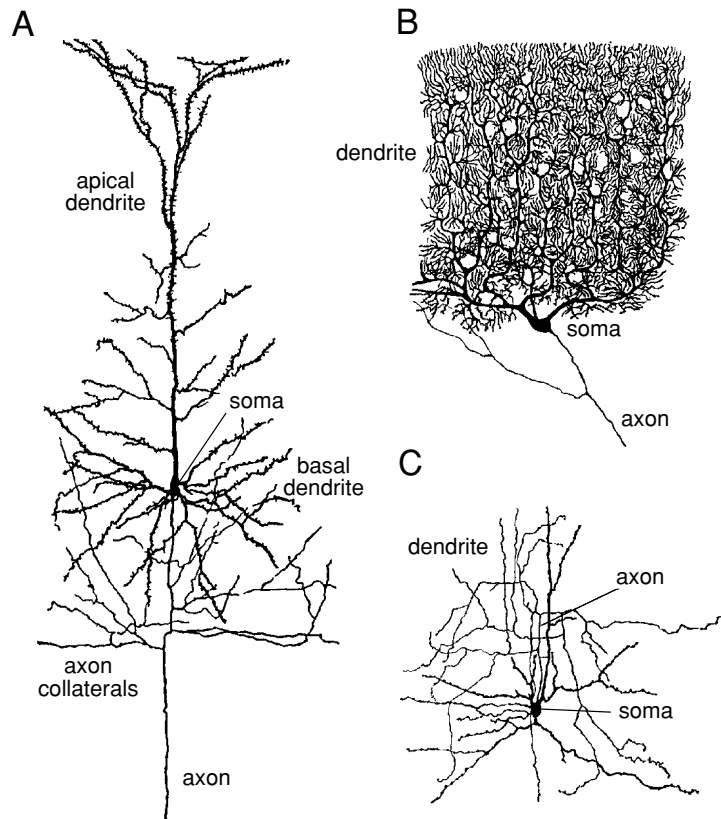
Figure 1.1 Diagrams of three neurons. (A) A cortical pyramidal cell. These are the primary excitatory neurons of the cerebral cortex. Pyramidal cell axons branch locally, sending axon collaterals to synapse with nearby neurons, and also project more distally to conduct signals to other parts of the brain and nervous system. (B) A Purkinje cell of the cerebellum. Purkinje cell axons transmit the output of the cerebellar cortex. (C) A stellate cell of the cerebral cortex. Stellate cells are one of a large class of interneurons that provide inhibitory input to the neurons of the cerebral cortex. These figures are magnified about 150-fold. (Drawings from Cajal, 1911; figure from Dowling, 1992.)

Action potentials are of great importance because they are the only form of membrane potential fluctuation that can propagate over large distances. Subthreshold potential fluctuations are severely attenuated over distances of 1 mm or less. Action potentials, on the other hand, are regenerated actively along axon processes and can travel rapidly over large distances without attenuation.

Axons terminate at synapses where the voltage transient of the action potential opens ion channels, producing an influx of Ca^{2+} that leads to the release of a neurotransmitter (figure 1.2B). The neurotransmitter binds to receptors at the signal-receiving or postsynaptic side of the synapse,
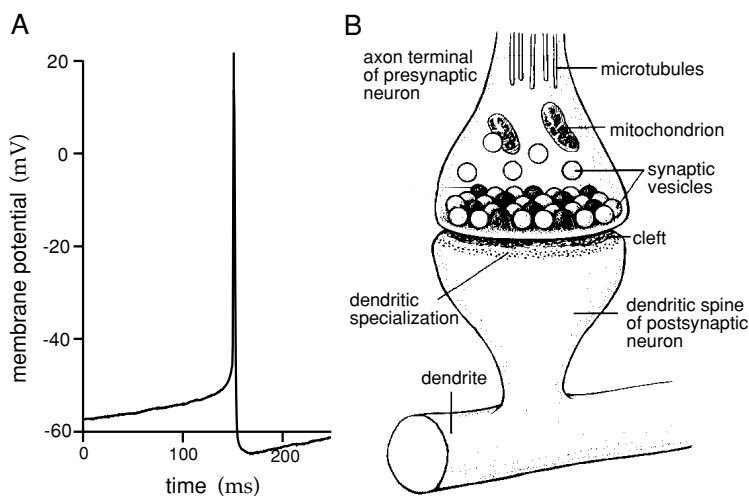
synapse

Figure 1.2 (A) An action potential recorded intracellularly from a cultured rat neocortical pyramidal cell. (B) Diagram of a synapse. The axon terminal or bouton is at the end of the axonal branch seen entering from the top of the figure. It is filled with synaptic vesicles containing the neurotransmitter that is released when an action potential arrives from the presynaptic neuron. Transmitter crosses the synaptic cleft and binds to receptors on the dendritic spine, a process roughly 1 μm long that extends from the dendrite of the postsynaptic neuron. Excitatory synapses onto cortical pyramidal cells form on dendritic spines as shown here. Other synapses form directly on the dendrites, axon, or soma of the postsynaptic neuron. (A recorded by L. Rutherford in the laboratory of G. Turrigiano; B adapted from Kandel et al., 1991.)

causing ion-conducting channels to open. Depending on the nature of the ion flow, the synapses can have either an excitatory, depolarizing, or an inhibitory, typically hyperpolarizing, effect on the postsynaptic neuron.

Recording Neuronal Responses

Figure 1.3 illustrates intracellular and extracellular methods for recording neuronal responses electrically (they can also be recorded optically). Membrane potentials are measured intracellularly by connecting a hollow glass electrode filled with a conducting electrolyte to a neuron, and comparing the potential it records with that of a reference electrode placed in the extracellular medium. Intracellular recordings are made either with

sharp and patch electrodes

sharp electrodes inserted through the membrane into the cell, or patch electrodes that have broader tips and are sealed tightly to the surface of the membrane. After the patch electrode seals, the membrane beneath its tip is either broken or perforated, providing electrical contact with the interior of the cell. The top trace in figure 1.3 is a schematic of an intracellular recording from the soma of a neuron firing a sequence of action potentials. The recording shows rapid spikes riding on top of a more slowly varying subthreshold potential. The bottom trace is a schematic of an intracellular recording made some distance out on the axon of the neuron. These traces
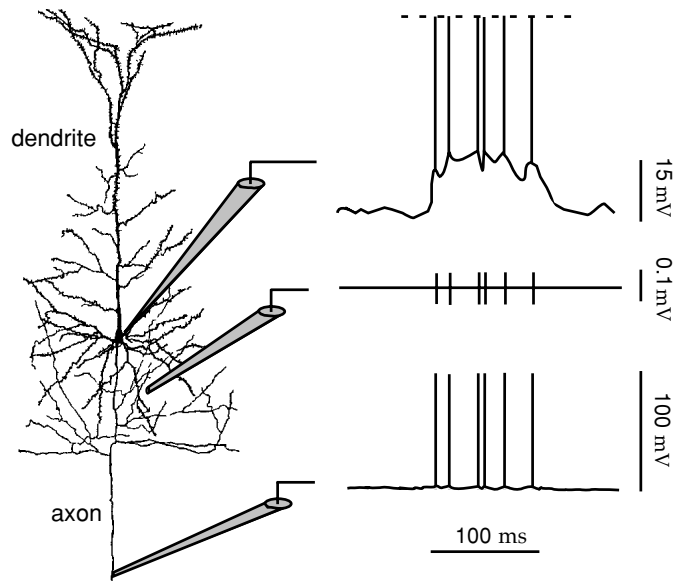
Figure 1.3 Three simulated recordings from a neuron. The top trace represents a recording from an intracellular electrode connected to the soma of the neuron. The height of the action potentials has been clipped to show the subthreshold membrane potential more clearly. The time scale is such that the action potential trajectory cannot be resolved. The bottom trace represents a recording from an intracellular electrode connected to the axon some distance away from the soma. The full height of the action potentials is indicated in this trace. The middle trace is a simulated extracellular recording. Action potentials appear as roughly equal positive and negative potential fluctuations with an amplitude of around 0.1 mV. This is roughly 1000 times smaller than the approximately 0.1 V amplitude of an intracellularly recorded action potential. (Neuron drawing is the same as figure 1.1A.)

are drawings, not real recordings; such intracellular axon recordings, although possible in some types of cells, are difficult and rare. Intracellular recordings from the soma are the norm, but intracellular dendritic recordings are increasingly being made as well. The subthreshold membrane potential waveform, apparent in the soma recording, is completely absent on the axon due to attenuation, but the action potential sequence in the two recordings is the same. This illustrates the important point that spikes, but not subthreshold potentials, propagate regeneratively down axons.

The middle trace in figure 1.3 illustrates an idealized, noise-free extracellular recording. Here an electrode is placed near a neuron but it does not penetrate the cell membrane. Such recordings can reveal the action potentials fired by a neuron, but not its subthreshold membrane potentials. Extracellular recordings are typically used for in vivo experiments, especially those involving behaving animals. Intracellular recordings are sometimes made in vivo, but this is difficult to do. Intracellular recording is more commonly used for in vitro preparations, such as slices of neural tissue. The responses studied in this chapter are action potential sequences that can be recorded either intra- or extracellularly.

extracellular electrodes

From Stimulus to Response

Characterizing the relationship between stimulus and response is difficult because neuronal responses are complex and variable. Neurons typically respond by producing complex spike sequences that reflect both the intrinsic dynamics of the neuron and the temporal characteristics of the stimulus. Isolating features of the response that encode changes in the stimulus can be difficult, especially if the time scale for these changes is of the same order as the average interval between spikes. Neural responses can vary from trial to trial even when the same stimulus is presented repeatedly. There are many potential sources of this variability, including variable levels of arousal and attention, randomness associated with various biophysical processes that affect neuronal firing, and the effects of other cognitive processes taking place during a trial. The complexity and trial-to-trial variability of action potential sequences make it unlikely that we can describe and predict the timing of each spike deterministically. Instead, we seek a model that can account for the probabilities that different spike sequences are evoked by a specific stimulus.

Typically, many neurons respond to a given stimulus, and stimulus features are therefore encoded by the activities of large neural populations. In studying population coding, we must examine not only the firing patterns of individual neurons but also the relationships of these firing patterns to each other across the population of responding cells.

In this chapter, we introduce the firing rate and spike-train correlation functions, which are basic measures of spiking probability and statistics. We also discuss spike-triggered averaging, a method for relating action potentials to the stimulus that evoked them. Finally, we present basic stochastic descriptions of spike generation, the homogeneous and inhomogeneous Poisson models, and discuss a simple model of neural responses to which they lead. In chapter 2, we continue our discussion of neural encoding by showing how reverse-correlation methods are used to construct estimates of firing rates in response to time-varying stimuli. These methods have been applied extensively to neural responses in the retina, lateral geniculate nucleus (LGN) of the thalamus, and primary visual cortex, and we review the resulting models.

1.2   Spike Trains and Firing Rates

Action potentials convey information through their timing. Although action potentials can vary somewhat in duration, amplitude, and shape, they are typically treated as identical stereotyped events in neural encoding studies. If we ignore the brief duration of an action potential (about 1 ms), an action potential sequence can be characterized simply by a list of the times when spikes occurred. For n spikes, we denote these times by t_i with $i = 1, 2, \ldots, n$. The trial during which the spikes are recorded is taken to

start at time 0 and end at time T, so $0 \leq t_i \leq T$ for all i. The spike sequence can also be represented as a sum of infinitesimally narrow, idealized spikes in the form of Dirac δ functions (see the Mathematical Appendix),

$$\rho(t) = \sum_{i=1}^{n} \delta(t - t_i) \,. \tag{1.1}$$

We call $\rho(t)$ the neural response function and use it to re-express sums over spikes as integrals over time. For example, for any well-behaved function $h(t)$, we can write

neural response function $\rho(t)$

$$\sum_{i=1}^{n} h(t - t_i) = \int_{-\infty}^{\infty} d\tau \, h(\tau) \rho(t - \tau) \,, \tag{1.2}$$

where the integral is over the duration of the trial. The equality follows from the basic defining equation for a δ function,

δ function

$$\int d\tau \, \delta(t - \tau) h(\tau) = h(t) \,, \tag{1.3}$$

provided that the limits of the integral surround the point t (if they do not, the integral is 0).

Because the sequence of action potentials generated by a given stimulus varies from trial to trial, neuronal responses are typically treated statistically or probabilistically. For example, they may be characterized by firing rates, rather than as specific spike sequences. Unfortunately, the term "firing rate" is applied conventionally to a number of different quantities. The simplest of these is what we call the spike-count rate, which is obtained by counting the number of action potentials that appear during a trial and dividing by the duration of the trial. We denote the spike-count rate by r, where

spike-count rate r

$$r = \frac{n}{T} = \frac{1}{T} \int_{0}^{T} d\tau \, \rho(\tau) \,. \tag{1.4}$$

The second equality follows from the fact that $\int d\tau \, \rho(\tau) = n$ and indicates that the spike-count rate is the time average of the neural response function over the duration of the trial.

The spike-count rate can be determined from a single trial, but at the expense of losing all temporal resolution about variations in the neural response during the course of the trial. A time-dependent firing rate can be defined by counting spikes over short time intervals, but this can no longer be computed from a single trial. For example, we can define the firing rate at time t during a trial by counting all the spikes that occurred between times t and $t + \Delta t$, for some small interval Δt, and dividing this count by Δt. However, for small Δt, which allows for high temporal resolution, the result of the spike count on any given trial is apt to be either 0 or 1, giving only two possible firing-rate values. The solution to this problem is to average over multiple trials. Thus, we define the time-dependent firing rate

as the average number of spikes (averaged over trials) appearing during a short interval between times t and $t + \Delta t$, divided by the duration of the interval.

The number of spikes occurring between times t and $t + \Delta t$ on a single trial is the integral of the neural response function over that time interval. The average number of spikes during this interval is the integral of the trial-averaged neural response function. We use angle brackets, $\langle\ \rangle$, to denote *trial average* $\langle\ \rangle$ averages over trials that use the same stimulus, so that $\langle z\rangle$ for any quantity z is the sum of the values of z obtained from many trials involving the same stimulus, divided by the number of trials. The trial-averaged neural response function is denoted by $\langle\rho(t)\rangle$, and the time-dependent firing rate *firing rate* $r(t)$ is given by

$$r(t) = \frac{1}{\Delta t} \int_{t}^{t+\Delta t} d\tau\, \langle\rho(\tau)\rangle\,. \tag{1.5}$$

We use the notation $r(t)$ for this important quantity (as opposed to r for the spike-count rate), and when we use the term "firing rate" without any modifiers, we mean $r(t)$. Formally, the limit $\Delta t \to 0$ should be taken on the right side of this expression, but, in extracting a time-dependent firing rate from data, the value of Δt must be large enough so there are sufficient numbers of spikes within the interval defining $r(t)$ to obtain a reliable estimate of the average.

For sufficiently small Δt, $r(t)\Delta t$ is the average number of spikes occurring between times t and $t + \Delta t$ over multiple trials. The average number of spikes over a longer time interval is given by the integral of $r(t)$ over that interval. If Δt is small, there will never be more than one spike within the interval between t and $t + \Delta t$ on any given trial. This means that $r(t)\Delta t$ is also the fraction of trials on which a spike occurred between those times. Equivalently, $r(t)\Delta t$ is the probability that a spike occurs during this time *spiking probability* interval. This probabilistic interpretation provides a formal definition of the time-dependent firing rate; $r(t)\Delta t$ is the probability of a spike occurring during a short interval of duration Δt around the time t.

In any integral expression such as equation 1.2, the neural response function generates a contribution whenever a spike occurs. If we use the trial-average response function instead, as in equation 1.5, this generates contributions proportional to the fraction of trials on which a spike occurred. Because of the relationship between this fraction and the firing rate, we can replace the trial-averaged neural response function with the firing rate $r(t)$ within any well-behaved integral, for example,

$$\int d\tau\, h(\tau)\, \langle\rho(t-\tau)\rangle = \int d\tau\, h(\tau) r(t-\tau) \tag{1.6}$$

for any function h. This establishes an important relationship between the average neural response function and the firing rate; the two are equivalent when used inside integrals. It also provides another interpretation of $r(t)$ as the trial-averaged density of spikes along the time axis.

In the same way that the response function $\rho(t)$ can be averaged across trials to give the firing rate $r(t)$, the spike-count firing rate can be averaged over trials, yielding a quantity that we refer to as the average firing rate. This is denoted by $\langle r \rangle$ and is given by

$$\langle r \rangle = \frac{\langle n \rangle}{T} = \frac{1}{T} \int_0^T d\tau \, \langle \rho(\tau) \rangle = \frac{1}{T} \int_0^T dt \, r(t) \,. \qquad (1.7)$$

The first equality indicates that $\langle r \rangle$ is just the average number of spikes per trial divided by the trial duration. The third equality follows from the equivalence of the firing rate and the trial-averaged neural response function within integrals (equation 1.6). The average firing rate is equal to both the time average of $r(t)$ and the trial average of the spike-count rate r. Of course, a spike-count rate and average firing rate can be defined by counting spikes over any time period, not necessarily the entire duration of a trial.

The term "firing rate" is commonly used for all three quantities, $r(t)$, r, and $\langle r \rangle$. Whenever possible, we use the terms "firing rate", "spike-count rate", and "average firing rate" for $r(t)$, r, and $\langle r \rangle$, respectively, but when this becomes too cumbersome, the different mathematical notations serve to distinguish them. In particular, we distinguish the spike-count rate r from the time-dependent firing rate $r(t)$ by using a different font and by including the time argument in the latter expression (unless $r(t)$ is independent of time). The difference between the fonts is rather subtle, but the context should make it clear which rate is being used.

Measuring Firing Rates

The firing rate $r(t)$ cannot be determined exactly from the limited data available from a finite number of trials. In addition, there is no unique way to approximate $r(t)$. A discussion of the different methods allows us to introduce the concept of a linear filter and kernel that will be used extensively in the following chapters. We illustrate these methods by extracting firing rates from a single trial, but more accurate results could be obtained by averaging over multiple trials.

Figure 1.4 compares a number of ways of approximating $r(t)$ from a spike sequence. Figure 1.4A shows 3 s of the response of a neuron in the inferotemporal cortex recorded while a monkey watched a video. Neurons in the region of cortex where this recording was made are selective for complex visual images, including faces. A simple way of extracting an estimate of the firing rate from a spike train like this is to divide time into discrete bins of duration Δt, count the number of spikes within each bin, and divide by Δt. Figure 1.4B shows the approximate firing rate computed using this procedure with a bin size of 100 ms. Note that with this procedure, the quantity being computed is really the spike-count firing rate over the duration of the bin, and that the firing rate $r(t)$ within a given bin is approximated by this spike-count rate.
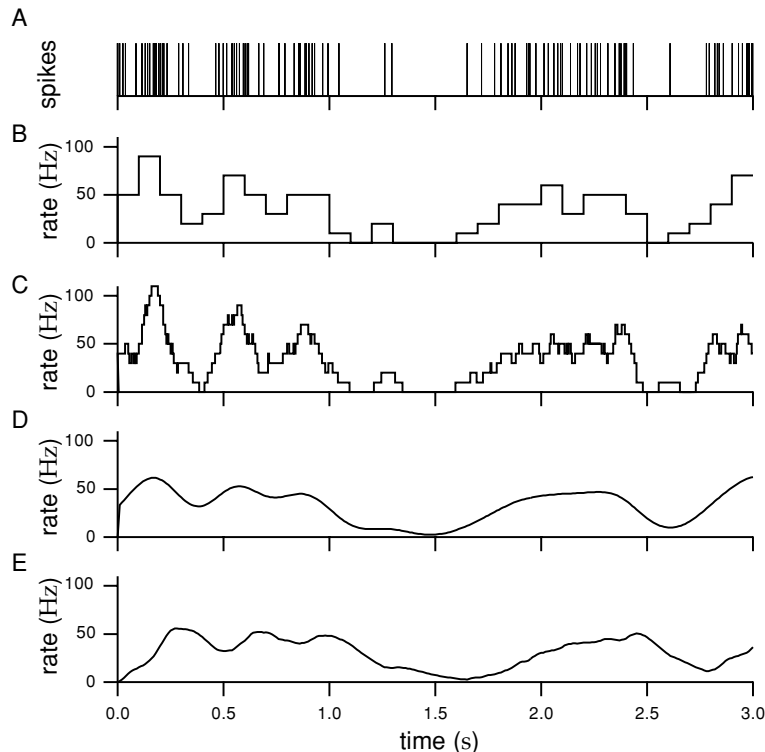
Figure 1.4 Firing rates approximated by different procedures. (A) A spike train from a neuron in the inferotemporal cortex of a monkey recorded while that animal watched a video on a monitor under free viewing conditions. (B) Discrete-time firing rate obtained by binning time and counting spikes with $\Delta t = 100$ ms. (C) Approximate firing rate determined by sliding a rectangular window function along the spike train with $\Delta t = 100$ ms. (D) Approximate firing rate computed using a Gaussian window function with $\sigma_t = 100$ ms. (E) Approximate firing rate using the window function of equation 1.12 with $1/\alpha = 100$ ms. (Data from Baddeley et al., 1997.)

The binning and counting procedure illustrated in figure 1.4B generates an estimate of the firing rate that is a piecewise constant function of time, resembling a histogram. Because spike counts can take only integer values, the rates computed by this method will always be integer multiples of $1/\Delta t$, and thus they take discrete values. Decreasing the value of Δt increases temporal resolution by providing an estimate of the firing rate at more finely spaced intervals of time, but at the expense of decreasing the resolution for distinguishing different rates. One way to avoid quantized firing rates is to vary the bin size so that a fixed number of spikes appears in each bin. The firing rate is then approximated as that fixed number of spikes divided by the variable bin width.

Counting spikes in preassigned bins produces a firing-rate estimate that depends not only on the size of the time bins but also on their placement. To avoid the arbitrariness in the placement of bins, we can instead take a single bin or window of duration Δt and slide it along the spike train,

counting the number of spikes within the window at each location. The jagged curve in figure 1.4C shows the result of sliding a 100 ms wide window along the spike train. The firing rate approximated in this way can be expressed as the sum of a window function over the times t_i for $i = 1, 2, \ldots, n$ when the n spikes in a particular sequence occurred,

$$r_{\text{approx}}(t) = \sum_{i=1}^{n} w(t - t_i)\,, \tag{1.8}$$

where the window function is

$$w(t) = \begin{cases} 1/\Delta t & \text{if } -\Delta t/2 \leq t < \Delta t/2 \\ 0 & \text{otherwise}\,. \end{cases} \tag{1.9}$$

Use of a sliding window avoids the arbitrariness of bin placement and produces a rate that might appear to have a better temporal resolution. However, it must be remembered that the rates obtained at times separated by less than one bin width are correlated because they involve some of the same spikes.

The sum in equation 1.8 can also be written as the integral of the window function times the neural response function (see equation 1.2):

$$r_{\text{approx}}(t) = \int_{-\infty}^{\infty} d\tau\, w(\tau)\rho(t - \tau)\,. \tag{1.10}$$

The integral in equation 1.10 is called a linear filter, and the window function w, also called the filter kernel, specifies how the neural response function evaluated at time $t - \tau$ contributes to the firing rate approximated at time t.

linear filter and kernel

The jagged appearance of the curve in figure 1.4C is caused by the discontinuous shape of the window function used. An approximate firing rate can be computed using virtually any window function $w(\tau)$ that goes to 0 outside a region near $\tau = 0$, provided that its time integral is equal to 1. For example, instead of the rectangular window function used in figure 1.4C, $w(\tau)$ can be a Gaussian:

$$w(\tau) = \frac{1}{\sqrt{2\pi}\sigma_w} \exp\left(-\frac{\tau^2}{2\sigma_w^2}\right)\,. \tag{1.11}$$

In this case, σ_w controls the temporal resolution of the resulting rate, playing a role analogous to Δt. A continuous window function like the Gaussian used in equation 1.8 generates a firing-rate estimate that is a smooth function of time (figure 1.4D).

Both the rectangular and the Gaussian window functions approximate the firing rate at any time, using spikes fired both before and after that time. A postsynaptic neuron monitoring the spike train of a presynaptic cell has access only to spikes that have previously occurred. An approximation of the firing rate at time t that depends only on spikes fired before t can be calculated using a window function that vanishes when its argument

is negative. Such a window function or kernel is called causal. One commonly used form is the α function

$$w(\tau) = [\alpha^2 \tau \exp(-\alpha\tau)]_+ \tag{1.12}$$

half-wave rectification [ ]$_+$

where $1/\alpha$ determines the temporal resolution of the resulting firing-rate estimate. The notation $[z]_+$ for any quantity z stands for the half-wave rectification operation,

$$[z]_+ = \begin{cases} z & \text{if } z \geq 0 \\ 0 & \text{otherwise}. \end{cases} \tag{1.13}$$

Figure 1.4E shows the firing rate approximated by such a causal scheme. Note that this rate tends to peak later than the rate computed in figure 1.4D using a temporally symmetric window function.

Tuning Curves

stimulus s

Neuronal responses typically depend on many different properties of a stimulus. In this chapter, we characterize responses of neurons as functions of just one of the stimulus attributes to which they may be sensitive. The value of this single attribute is denoted by s. In chapter 2, we consider more complete stimulus characterizations.

response tuning curve $f(s)$

A simple way of characterizing the response of a neuron is to count the number of action potentials fired during the presentation of a stimulus. This approach is most appropriate if the parameter s characterizing the stimulus is held constant over the trial. If we average the number of action potentials fired over (in theory, an infinite number of) trials and divide by the trial duration, we obtain the average firing rate, $\langle r \rangle$, defined in equation 1.7. The average firing rate written as a function of s, $\langle r \rangle = f(s)$, is called the neural response tuning curve. The functional form of a tuning curve depends on the parameter s used to describe the stimulus. The precise choice of parameters used as arguments of tuning curve functions is partially a matter of convention. Because tuning curves correspond to firing rates, they are measured in units of spikes per second or Hz.

primary visual cortex V1

Figure 1.5A shows extracellular recordings of a neuron in the primary visual cortex (V1) of a monkey. While these recordings were being made, a bar of light was moved at different angles across the region of the visual field where the cell responded to light. This region is called the receptive field of the neuron. Note that the number of action potentials fired depends on the angle of orientation of the bar. The same effect is shown in figure 1.5B in the form of a response tuning curve, which indicates how the average firing rate depends on the orientation of the light bar stimulus.

Gaussian tuning curve

The data have been fitted by a response tuning curve of the form

$$f(s) = r_{\max} \exp\left(-\frac{1}{2}\left(\frac{s - s_{\max}}{\sigma_f}\right)^2\right), \tag{1.14}$$
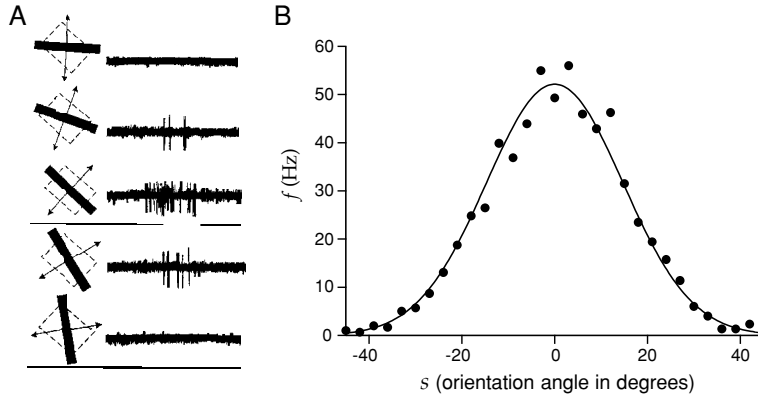
Figure 1.5 (A) Recordings from a neuron in the primary visual cortex of a monkey. A bar of light was moved across the receptive field of the cell at different angles. The diagrams to the left of each trace show the receptive field as a dashed square and the light source as a black bar. The bidirectional motion of the light bar is indicated by the arrows. The angle of the bar indicates the orientation of the light bar for the corresponding trace. (B) Average firing rate of a cat V1 neuron plotted as a function of the orientation angle of the light bar stimulus. The curve is a fit using the function 1.14 with parameters $r_{max} = 52.14$ Hz, $s_{max} = 0°$, and $\sigma_f = 14.73°$. (A adapted from Wandell, 1995, based on an original figure from Hubel and Wiesel, 1968; B data points from Henry et al., 1974).)

where s is the orientation angle of the light bar, s_{max} is the orientation angle evoking the maximum average response rate r_{max} (with $s - s_{max}$ taken to lie in the range between $-90°$ and $+90°$), and σ_f determines the width of the tuning curve. The neuron responds most vigorously when a stimulus having $s = s_{max}$ is presented, so we call s_{max} the preferred orientation angle of the neuron.

Response tuning curves can be used to characterize the selectivities of neurons in visual and other sensory areas to a variety of stimulus parameters. Tuning curves can also be measured for neurons in motor areas, in which case the average firing rate is expressed as a function of one or more parameters describing a motor action. Figure 1.6A shows an example of extracellular recordings from a neuron in primary motor cortex in a monkey *primary motor* that has been trained to reach in different directions. The stacked traces for *cortex M1* each direction are rasters showing the results of five different trials. The horizontal axis in these traces represents time, and each mark indicates an action potential. The firing pattern of the cell, in particular the rate at which spikes are generated, is correlated with the direction of arm movement, and thus encodes information about this aspect of the motor action.

Figure 1.6B shows the response tuning curve of an M1 neuron plotted as *cosine* a function of the direction of arm movement. Here the data points have *tuning curve* been fitted by a tuning curve of the form

$$f(s) = r_0 + (r_{max} - r_0)\cos(s - s_{max}), \qquad (1.15)$$

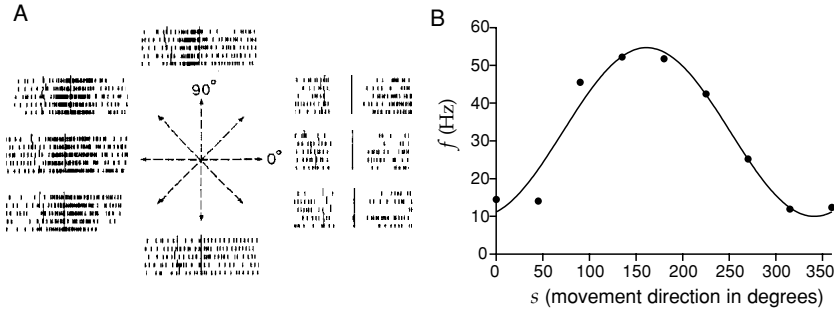where s is the reaching angle of the arm, s_{max} is the reaching angle associ-

Figure 1.6 (A) Recordings from the primary motor cortex of a monkey performing an arm-reaching task. The hand of the monkey started from a central resting location, and reaching movements were made in the directions indicated by the arrows. The rasters for each direction show action potentials fired on five trials. (B) Average firing rate plotted as a function of the direction in which the monkey moved its arm. The curve is a fit using the function 1.15 with parameters $r_{max} = 54.69$ Hz, $r_0 = 32.34$ Hz, and $s_{max} = 161.25°$. (A adapted from Georgopoulos et al., 1982, which is also the source of the data points in B.)

ated with the maximum response r_{max}, and r_0 is an offset or background firing rate that shifts the tuning curve up from the zero axis. The minimum firing rate predicted by equation 1.15 is $2r_0 - r_{max}$. For the neuron of figure 1.6B, this is a positive quantity, but for some M1 neurons $2r_0 - r_{max} < 0$, and the function 1.15 is negative over some range of angles. Because firing rates cannot be negative, the cosine tuning curve must be half-wave rectified in these cases (see equation 1.13),

$$f(s) = [r_0 + (r_{max} - r_0)\cos(s - s_{max})]_+ .$$ (1.16)

Figure 1.7B shows how the average firing rate of a V1 neuron depends on retinal disparity and illustrates another important type of tuning curve. Retinal disparity is a difference in the retinal location of an image between the two eyes (figure 1.7A). Some neurons in area V1 are sensitive to disparity, representing an early stage in the representation of viewing distance. In figure 1.7B, the data points have been fitted with a tuning curve called *sigmoidal tuning curve* a logistic or sigmoidal function,

$$f(s) = \frac{r_{max}}{1 + \exp\left((s_{1/2} - s)/\Delta_s\right)} .$$ (1.17)

In this case, s is the retinal disparity, the parameter $s_{1/2}$ is the disparity that produces a firing rate half as big as the maximum value r_{max}, and Δ_s controls how quickly the firing rate increases as a function of s. If Δ_s is negative, the firing rate is a monotonically decreasing function of s rather than a monotonically increasing function as in figure 1.7B.

Spike-Count Variability

Tuning curves allow us to predict the average firing rate, but they do not describe how the spike-count firing rate r varies about its mean value
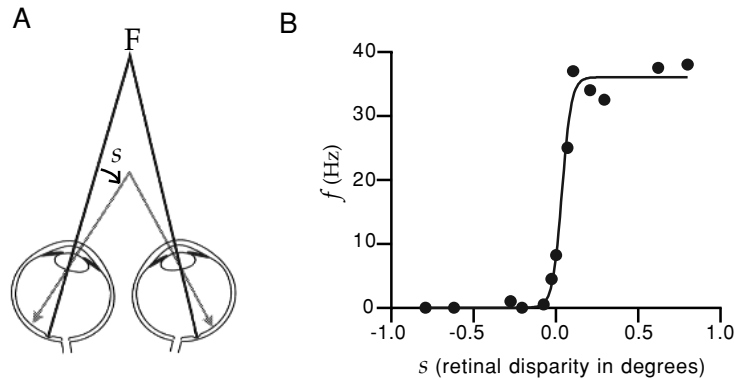
Figure 1.7 (A) Definition of retinal disparity. The gray lines with arrows show the location on each retina of an object located nearer than the fixation point F. The image from the fixation point falls at the fovea in each eye, the small pit where the black lines meet the retina. The image from a nearer object falls to the left of the fovea in the left eye and to the right of the fovea in the right eye. For objects farther away than the fixation point, this would be reversed. The disparity angle s is indicated in the figure. (B) Average firing rate of a cat V1 neuron responding to separate bars of light illuminating each eye, plotted as a function of the disparity. Because this neuron fires for positive s values, it is called a far-tuned cell. The curve is a fit using the function 1.17 with parameters $r_{max} = 36.03$ Hz, $s_{1/2} = 0.036°$, and $\Delta_s = 0.029°$. (A adapted from Wandell, 1995; B data points from Poggio and Talbot, 1981.)

$\langle r \rangle = f(s)$ from trial to trial. While the map from stimulus to average response may be described deterministically, it is likely that single-trial responses such as spike-count rates can be modeled only in a probabilistic manner. For example, r values can be generated from a probability distribution with mean $f(s)$. The trial-to-trial deviation of r from $f(s)$ is considered to be noise, and such models are often called noise models. The standard deviation for the noise distribution either can be independent of $f(s)$, in which case the variability is called additive noise, or it can depend on $f(s)$. Multiplicative noise corresponds to having the standard deviation proportional to $f(s)$.

Response variability extends beyond the level of spike counts to the entire temporal pattern of action potentials. Later in this chapter, we discuss a model of the neuronal response that uses a stochastic spike generator to produce response variability. This approach takes a deterministic estimate of the firing rate, $r_{est}(t)$, and produces a stochastic spiking pattern from it. The spike generator produces variable numbers and patterns of action potentials, even if the same estimated firing rate is used on each trial.

1.3 What Makes a Neuron Fire?

Response tuning curves characterize the average response of a neuron to a given stimulus. We now consider the complementary procedure of av-

eraging the stimuli that produce a given response. To average stimuli in this way, we need to specify what fixed response we will use to "trigger" the average. The most obvious choice is the firing of an action potential. Thus, we ask, "What, on average, did the stimulus do before an action potential was fired?" The resulting quantity, called the spike-triggered average stimulus, provides a useful way of characterizing neuronal selectivity. Spike-triggered averages are computed using stimuli characterized by a parameter $s(t)$ that varies over time. Before beginning our discussion of spike triggering, we describe some features of such stimuli.

Describing the Stimulus

Neurons responding to sensory stimuli face the difficult task of encoding parameters that can vary over an enormous dynamic range. For example, photoreceptors in the retina can respond to single photons or can operate in bright light with an influx of millions of photons per second. To deal with such wide-ranging stimuli, sensory neurons often respond most strongly to rapid changes in stimulus properties and are relatively insensitive to steady-state levels. Steady-state responses are highly compressed functions of stimulus intensity, typically with logarithmic or weak power-law dependences. This compression has an interesting psychophysical correlate. Weber measured how different the intensity of two stimuli had to be for them to be reliably discriminated, the "just noticeable" difference Δs. He found that, for a given stimulus, Δs is proportional to the magnitude of the stimulus s, so that $\Delta s/s$ is constant. This relationship is called *Weber's law*. Fechner suggested that noticeable differences set the scale for perceived stimulus intensities. Integrating Weber's law, this means that the perceived intensity of a stimulus of absolute intensity s varies as *Fechner's law* $\log s$. This is known as Fechner's law.

Weber's law

Fechner's law

Sensory systems make numerous adaptations, using a variety of mechanisms, to adjust to the average level of stimulus intensity. When a stimulus generates such adaptation, the relationship between stimulus and response is often studied in a potentially simpler regime by describing responses to fluctuations about a mean stimulus level. In this case, $\int_0^T dt\, s(t)/T = 0$ $s(t)$ is defined so that its time average over the duration of a trial is 0, $\int_0^T dt\, s(t)/T = 0$. We frequently impose this condition.

Our analysis of neural encoding involves two different types of averages: averages over repeated trials that employ the same stimulus, which we denote by angle brackets, and averages over different stimuli. We could introduce a second notation for averages over stimuli, but this can be avoided when using time-dependent stimuli. Instead of presenting a number of different stimuli and averaging over them, we can string together all of the stimuli we wish to consider into a single time-dependent stimulus *stimulus and* sequence and average over time. Thus, stimulus averages are replaced by *time averages* time averages.

Although a response recorded over a trial depends only on the values

taken by $s(t)$ during that trial, some of the mathematical analyses presented in this chapter and in chapter 2 are simplified if we define the stimulus at other times as well. It is convenient if integrals involving the stimulus are time-translationally invariant so that for any function h and time interval τ

$$\int_0^T dt\, h(s(t+\tau)) = \int_\tau^{T+\tau} dt\, h(s(t)) = \int_0^T dt\, h(s(t)).\qquad(1.18)$$

To assure the last equality, we define the stimulus outside the time limits of the trial by the relation $s(T+\tau) = s(\tau)$ for any τ, thereby making the stimulus periodic.   *periodic stimulus*

The Spike-Triggered Average

The spike-triggered average stimulus, $C(\tau)$, is the average value of the stimulus a time interval τ before a spike is fired. In other words, for a spike occurring at time t_i, we determine $s(t_i - \tau)$, and then we sum over all n spikes in a trial, $i = 1, 2, \dots, n$, and divide the total by n. In addition, we average over trials. Thus,   *spike-triggered average $C(\tau)$*

$$C(\tau) = \left\langle \frac{1}{n} \sum_{i=1}^{n} s(t_i - \tau) \right\rangle \approx \frac{1}{\langle n \rangle} \left\langle \sum_{i=1}^{n} s(t_i - \tau) \right\rangle.\qquad(1.19)$$

The approximate equality of the last expression follows from the fact that if n is large, the total number of spikes on each trial is well approximated by the average number of spikes per trial, $n \approx \langle n \rangle$. We make use of this approximation because it allows us to relate the spike-triggered average to other quantities commonly used to characterize the relationship between stimulus and response (see below). Figure 1.8 provides a schematic description of the computation of the spike-triggered average. Each time a spike appears, the stimulus in a time window preceding the spike is recorded. Although the range of τ values in equation 1.19 is unlimited, the response is typically affected only by the stimulus in a window a few hundred milliseconds wide immediately preceding a spike. More precisely, we expect $C(\tau)$ to approach 0 for positive τ values larger than the correlation time between the stimulus and the response. If the stimulus has no temporal correlations with itself, we also expect $C(\tau)$ to be 0 for $\tau < 0$, because the response of a neuron cannot depend on future stimuli. In practice, the stimulus is recorded only over a finite time period, as indicated by the shaded areas in figure 1.8. The recorded stimuli for all spikes are then summed and the procedure is repeated over multiple trials.

The spike-triggered average stimulus can be expressed as an integral of the stimulus times the neural response function of equation 1.1. If we replace the sum over spikes with an integral, as in equation 1.2, and use the approximate expression for $C(\tau)$ in equation 1.19, we find

$$C(\tau) = \frac{1}{\langle n \rangle} \int_0^T dt\, \langle \rho(t) \rangle\, s(t - \tau) = \frac{1}{\langle n \rangle} \int_0^T dt\, r(t)s(t - \tau).\qquad(1.20)$$
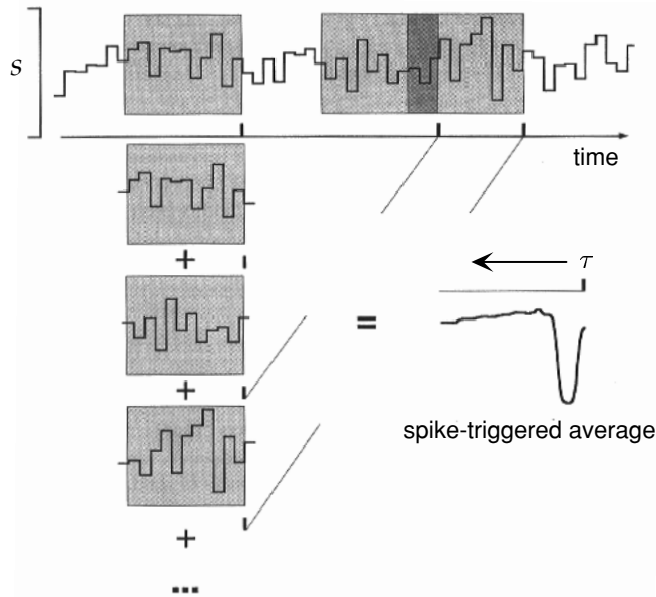
Figure 1.8 Schematic of the procedure for computing the spike-triggered average stimulus. Each gray rectangle contains the stimulus prior to one of the spikes shown along the time axis. These are averaged to produce the waveform shown at the lower right, which is the average stimulus before a spike. The stimulus in this example is a piecewise constant function of time. (Adapted from Rieke et al., 1997.)

The second equality is due to the equivalence of $\langle \rho(t) \rangle$ and $r(t)$ within integrals. Equation 1.20 allows us to relate the spike-triggered average to the correlation function of the firing rate and the stimulus.

Correlation functions are a useful way of determining how two quantities that vary over time are related to one another. The two quantities being related are evaluated at different times, one at time t and the other at time $t + \tau$. The correlation function is then obtained by averaging their product over all t values, and it is a function of τ. The correlation function of the firing rate and the stimulus is

firing-rate stimulus correlation function Q_{rs}

$$Q_{rs}(\tau) = \frac{1}{T} \int_0^T dt\, r(t) s(t + \tau)\,. \qquad (1.21)$$

By comparing equations 1.20 and 1.21, we find that

$$C(\tau) = \frac{1}{\langle r \rangle} Q_{rs}(-\tau)\,, \qquad (1.22)$$

where $\langle r \rangle = \langle n \rangle / T$ is the average firing rate over the set of trials. Because the argument of the correlation function in equation 1.22 is $-\tau$, the spike-triggered average stimulus is often called the reverse correlation function. It is proportional to the correlation of the firing rate with the stimulus at preceding times.
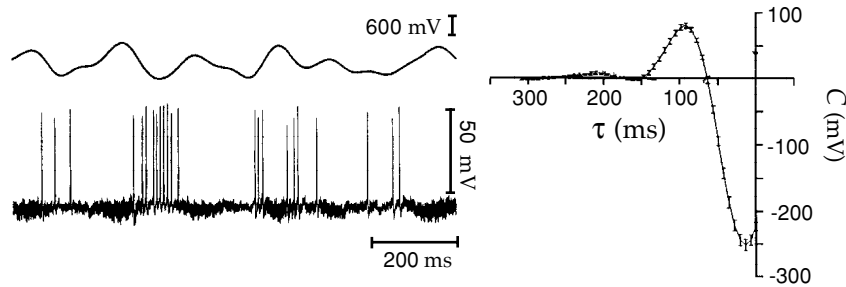
reverse correlation function

Figure 1.9 The spike-triggered average stimulus for a neuron of the electrosensory lateral-line lobe of the weakly electric fish *Eigenmannia*. The upper left trace is the potential used to generate the electric field to which this neuron is sensitive. The evoked spike train is plotted below the stimulus potential. The plot on the right is the spike-triggered average stimulus. (Adapted from Gabbiani et al., 1996.)

The spike-triggered average stimulus is widely used to study and characterize neural responses. Because $C(\tau)$ is the average value of the stimulus at a time τ before a spike, larger values of τ represent times farther in the past relative to the time of the triggering spike. For this reason, we plot spike-triggered averages with the time axis going backward compared to the normal convention. This allows the average spike-triggering stimulus to be read off from the plots in the usual left-to-right order.

Figure 1.9 shows the spike-triggered average stimulus for a neuron in the electrosensory lateral-line lobe of the weakly electric fish *Eigenmannia*. Weakly electric fish generate oscillating electric fields from an internal electric organ. Distortions in the electric field produced by nearby objects are detected by sensors spread over the skin of the fish. The lateral-line lobe acts as a relay station along the processing pathway for electrosensory signals. Fluctuating electrical potentials, such as that shown in the upper left trace of figure 1.9, elicit responses from electrosensory lateral-line lobe neurons, as seen in the lower left trace. The spike-triggered average stimulus, plotted at the right, indicates that, on average, the electric potential made a positive upswing followed by a large negative deviation prior to a spike being fired by this neuron.

The results obtained by spike-triggered averaging depend on the particular set of stimuli used during an experiment. How should this set be chosen? In chapter 2, we show that there are certain advantages to using a stimulus that is uncorrelated from one time to the next, a white-noise stimulus. A heuristic argument supporting the use of such stimuli is that in asking what makes a neuron fire, we may want to sample its responses to stimulus fluctuations at all frequencies with equal weight (i.e., equal power), and this is one of the properties of white-noise stimuli. In practice, white-noise stimuli can be generated with equal power only up to a finite frequency cutoff, but neurons respond to stimulus fluctuations only within a limited frequency range anyway. Figure 1.9 is based on such an approximate white-noise stimulus. The power in a signal as a function

of its frequency is called the power spectrum or power spectral density. White noise has a flat power spectrum.

White-Noise Stimuli

The defining characteristic of a white-noise stimulus is that its value at any one time is uncorrelated with its value at any other time. This condition can be expressed using the stimulus-stimulus correlation function, also called the stimulus autocorrelation, which is defined by analogy with equation 1.21 as

stimulus autocorrelation function Q_{ss}

$$Q_{ss}(\tau) = \frac{1}{T} \int_0^T dt\, s(t)s(t+\tau)\,. \tag{1.23}$$

Just as a correlation function provides information about the temporal relationship between two quantities, so an autocorrelation function tells us about how a quantity at one time is related to itself evaluated at another time. For white noise, the stimulus autocorrelation function is 0 in the range $-T/2 < \tau < T/2$ except when $\tau = 0$, and over this range

$$Q_{ss}(\tau) = \sigma_s^2 \delta(\tau)\,. \tag{1.24}$$

The constant σ_s, which has the units of the stimulus times the square root of the unit of time, reflects the magnitude of the variability of the white noise. In appendix A, we show that equation 1.24 is equivalent to the statement that white noise has equal power at all frequencies.

No physical system can generate noise that is white to arbitrarily high frequencies. Approximations of white noise that are missing high-frequency components can be used, provided the missing frequencies are well above the sensitivity of the neuron under investigation. To approximate white noise, we consider times that are integer multiples of a basic unit of duration Δt, that is, times $t = m\Delta t$ for $m = 1, 2, \dots, M$ where $M\Delta t = T$. The function $s(t)$ is then constructed as a discrete sequence of stimulus values. This produces a steplike stimulus waveform, like the one that appears in figure 1.8, with a constant stimulus value s_m presented during time bin m. In terms of the discrete-time values s_m, the condition that the stimulus is uncorrelated is

$$\frac{1}{M} \sum_{m=1}^{M} s_m s_{m+p} = \begin{cases} \sigma_s^2/\Delta t & \text{if} \quad p = 0 \\ 0 & \text{otherwise}\,. \end{cases} \tag{1.25}$$

The factor of $1/\Delta t$ on the right side of this equation reproduces the δ function of equation 1.24 in the limit $\Delta t \to 0$. For approximate white noise, the autocorrelation function is 0 except for a region around $\tau = 0$ with width of order Δt. Similarly, the binning of time into discrete intervals of size Δt means that the noise generated has a flat power spectrum only up to frequencies of order $1/(2\Delta t)$.
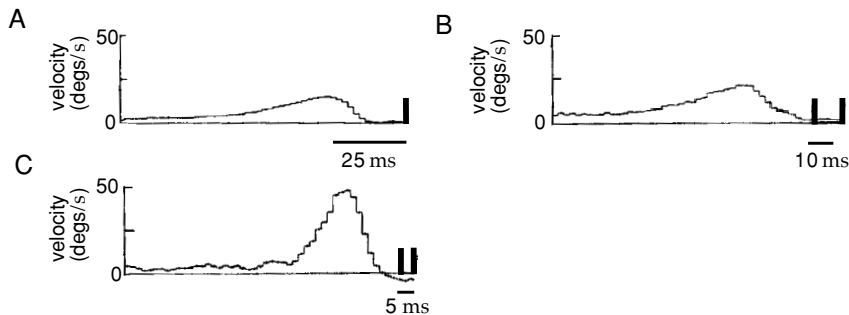
Figure 1.10 Single- and multiple-spike-triggered average stimuli for a blowfly H1 neuron responding to a moving visual image. (A) The average stimulus velocity triggered on a single spike. (B) The average stimulus velocity before two spikes with a separation of 10 ± 1 ms. (C) The average stimulus before two spikes with a separation of 5 ± 1 ms. (Data from de Ruyter van Steveninck and Bialek, 1988; figure adapted from Rieke et al., 1997.)

An approximation to white noise can be generated by choosing each s_m independently from a probability distribution with mean 0 and variance $\sigma_s^2/\Delta t$. Any reasonable probability function satisfying these two conditions can be used to generate the stimulus values within each time bin. A special class of white-noise stimuli, Gaussian white noise, results when the probability distribution used to generate the s_m values is a Gaussian function. The factor of $1/\Delta t$ in the variance indicates that the variability must be increased as the time bins get smaller. A number of other schemes for efficiently generating approximations of white-noise stimuli are discussed in the references at the end of this chapter.

Multiple-Spike-Triggered Averages and Spike-Triggered Correlations

In addition to triggering on single spikes, stimulus averages can be computed by triggering on various combinations of spikes. Figure 1.10 shows some examples of two-spike triggers. These results come from a study of the H1 movement-sensitive visual neuron of the blowfly. The H1 neuron detects the motion of visual images during flight in order to generate and guide stabilizing motor corrections. It responds to motion of the visual scene. In the experiments, the fly is held fixed while a visual image with a time-varying velocity $s(t)$ is presented. Figure 1.10A, showing the spike-triggered average stimulus, indicates that this neuron responds to positive angular velocities after a latency of about 15 ms. Figure 1.10B is the average stimulus prior to the appearance of two spikes separated by 10 ± 1 ms. In this case, the two-spike average is similar to the sum of two single-spike-triggered average stimuli displaced from one another by 10 ms. Thus, for 10 ms separations, two spikes occurring together tell us no more as a two-spike unit than they would individually. This result changes when shorter separations are considered. Figure 1.10C shows the

average stimulus triggered on two spikes separated by 5 ± 1 ms. The average stimulus triggered on a pair of spikes separated by 5 ms is not the same as the sum of the average stimuli for each spike separately.

Spike-triggered averages of other stimulus-dependent quantities can provide additional insight into neural encoding, for example, spike-triggered average autocorrelation functions. Obviously, spike-triggered averages of higher-order stimulus combinations can be considered as well.

1.4   Spike-Train Statistics

A complete description of the stochastic relationship between a stimulus and a response would require us to know the probabilities corresponding to every sequence of spikes that can be evoked by the stimulus. Spike times are continuous variables, and, as a result, the probability for a spike to occur at any precisely specified time is actually zero. To get a nonzero value, we must ask for the probability that a spike occurs within a specified interval, for example, the interval between times t and $t + \Delta t$. For small Δt, the probability of a spike falling in this interval is proportional to the size of the interval, Δt. A similar relation holds for any continuous stochastic variable z. The probability that z takes a value between z and $z + \Delta z$, for small Δz (strictly speaking, as $\Delta z \to 0$), is equal to $p[z]\Delta z$, where $p[z]$ is called a probability density.

Throughout this book, we use the notation $P[\,]$ to denote probabilities and $p[\,]$ to denote probability densities. We use the bracket notation $P[\,]$ generically for the probability of something occurring and also to denote a specific probability function. In the latter case, the notation $P(\,)$ would be more appropriate, but switching between square brackets and parentheses is confusing, so the reader will have to use the context to distinguish between these cases.

The probability of a spike sequence appearing is proportional to the probability density of spike times, $p[t_1, t_2, \ldots, t_n]$. In other words, the probability $P[t_1, t_2, \ldots, t_n]$ that a sequence of n spikes occurs with spike i falling between times t_i and $t_i + \Delta t$ for $i = 1, 2, \ldots, n$ is given in terms of this density by the relation $P[t_1, t_2, \ldots, t_n] = p[t_1, t_2, \ldots, t_n](\Delta t)^n$.

Unfortunately, the number of possible spike sequences is typically so large that determining or even roughly estimating all of their probabilities of occurrence is impossible. Instead, we must rely on some statistical model that allows us to estimate the probability of an arbitrary spike sequence occurring, given our knowledge of the responses actually recorded. The firing rate $r(t)$ determines the probability of firing a single spike in a small interval around the time t, but $r(t)$ is not, in general, sufficient information to predict the probabilities of spike sequences. For example, the probability of two spikes occurring together in a sequence is not necessarily equal to the product of the probabilities that they occur individually, because

the presence of one spike may effect the occurrence of the other. If, however, the probability of generating an action potential is independent of the presence or timing of other spikes (i.e., if the spikes are statistically independent) the firing rate is all that is needed to compute the probabilities for all possible action potential sequences.

A stochastic process that generates a sequence of events, such as action   *point process*
potentials, is called a point process. In general, the probability of an event
occurring at any given time could depend on the entire history of preceding events. If this dependence extends only to the immediately preceding event, so that the intervals between successive events are independent, the point process is called a renewal process. If there is no dependence   *renewal process*
at all on preceding events, so that the events themselves are statistically independent, we have a Poisson process. The Poisson process provides an extremely useful approximation of stochastic neuronal firing. To make   *Poisson process*
the presentation easier to follow, we separate two cases, the homogeneous Poisson process, for which the firing rate is constant over time, and the inhomogeneous Poisson process, which involves a time-dependent firing rate.

The Homogeneous Poisson Process

We denote the firing rate for a homogeneous Poisson process by $r(t) = r$ because it is independent of time. When the firing rate is constant, the Poisson process generates every sequence of n spikes over a fixed time interval with equal probability. As a result, the probability $P[t_1, t_2, \ldots, t_n]$ can be expressed in terms of another probability function $P_T[n]$, which is the probability that an arbitrary sequence of exactly n spikes occurs within a trial of duration T. Assuming that the spike times are ordered so that $0 \leq t_1 \leq t_2 \leq \ldots \leq t_n \leq T$, the relationship is

$$P[t_1, t_2, \ldots, t_n] = n! P_T[n] \left(\frac{\Delta t}{T} \right)^n . \qquad (1.26)$$

This relationship is a special case of equation 1.37 derived below.

To compute $P_T[n]$, we divide the time T into M bins of size $\Delta t = T/M$. We can assume that Δt is small enough so that we never get two spikes within any one bin because, at the end of the calculation, we take the limit $\Delta t \to 0$. $P_T[n]$ is the product of three factors: the probability of generating n spikes within a specified set of the M bins, the probability of not generating spikes in the remaining $M - n$ bins, and a combinatorial factor equal to the number of ways of putting n spikes into M bins. The probability of a spike occurring in one specific bin is $r\Delta t$, and the probability of n spikes appearing in n specific bins is $(r\Delta t)^n$. Similarly, the probability of not having a spike in a given bin is $(1 - r\Delta t)$, so the probability of having the remaining $M - n$ bins without any spikes in them is $(1 - r\Delta t)^{M-n}$. Finally, the number of ways of putting n spikes into M bins is given by the
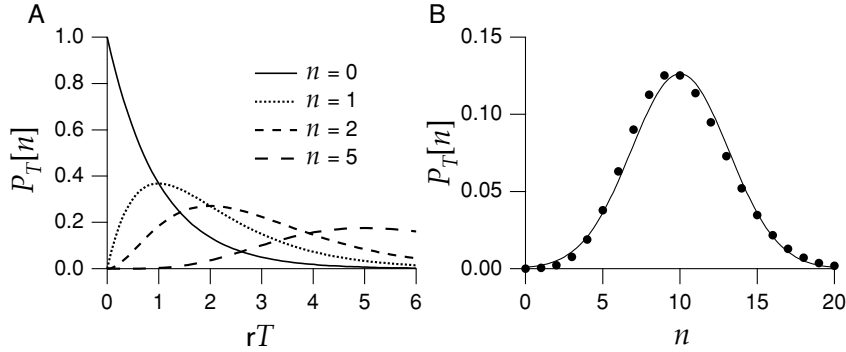
Figure 1.11 (A) The probability that a homogeneous Poisson process generates *n* spikes in a time period of duration *T* plotted for *n* = 0, 1, 2, and 5. The probability is plotted as function of the rate times the duration of the interval, r*T*, to make the plot applicable for any rate. (B) The probability of finding *n* spikes during a time period for which r*T* = 10 (dots) compared with a Gaussian distribution with mean and variance equal to 10 (line).

binomial coefficient $M!/(M-n)!n!$. Putting all these factors together, we find

$$P_T[n] = \lim_{\Delta t \to 0} \frac{M!}{(M-n)!n!}(r\Delta t)^n(1-r\Delta t)^{M-n}. \qquad (1.27)$$

To take the limit, we note that as $\Delta t \to 0$, M grows without bound because $M\Delta t = T$. Because n is fixed, we can write $M - n \approx M = T/\Delta t$. Using this approximation and defining $\epsilon = -r\Delta t$, we find that

$$\lim_{\Delta t \to 0}(1-r\Delta t)^{M-n} = \lim_{\epsilon \to 0}\left((1+\epsilon)^{1/\epsilon}\right)^{-rT} = e^{-rT} = \exp(-rT) \qquad (1.28)$$

because $\lim_{\epsilon \to 0}(1+\epsilon)^{1/\epsilon}$ is, by definition, $e = \exp(1)$. For large M, $M!/(M-n)! \approx M^n = (T/\Delta t)^n$, so

$$P_T[n] = \frac{(rT)^n}{n!}\exp(-rT). \qquad (1.29)$$

Poisson distribution This is called the Poisson distribution. The probabilities $P_T[n]$, for a few *n* values, are plotted as a function of r*T* in figure 1.11A. Note that as *n* increases, the probability reaches its maximum at larger *T* values and that large *n* values are more likely than small ones for large *T*. Figure 1.11B shows the probabilities of various numbers of spikes occurring when the average number of spikes is 10. For large r*T*, which corresponds to a large expected number of spikes, the Poisson distribution approaches a Gaussian distribution with mean and variance equal to r*T*. Figure 1.11B shows that this approximation is already quite good for r*T* = 10.

We can compute the variance of spike counts produced by a Poisson process from the probabilities in equation 1.29. For spikes counted over an interval of duration *T*, the variance of the spike count (derived in appendix B) is

$$\sigma_n^2 = \langle n^2 \rangle - \langle n \rangle^2 = rT. \qquad (1.30)$$

Thus the variance and mean of the spike count are equal. The ratio of these two quantities, $\sigma_n^2/\langle n \rangle$, is called the Fano factor and takes the value 1 for a homogeneous Poisson process, independent of the time interval T.

Fano factor

The probability density of time intervals between adjacent spikes is called the interspike interval distribution, and it is a useful statistic for characterizing spiking patterns. Suppose that a spike occurs at a time t_i for some value of i. The probability of a homogeneous Poisson process generating the next spike somewhere in the interval $t_i + \tau \leq t_{i+1} < t_i + \tau + \Delta t$, for small Δt, is the probability that no spike is fired for a time τ, times the probability, $r\Delta t$, of generating a spike within the following small interval Δt. From equation 1.29, with $n = 0$, the probability of not firing a spike for period τ is $\exp(-r\tau)$, so the probability of an interspike interval falling between τ and $\tau + \Delta t$ is

interspike interval distribution

$$P[\tau \leq t_{i+1} - t_i < \tau + \Delta t] = r\Delta t \exp(-r\tau) . \qquad (1.31)$$

The probability density of interspike intervals is, by definition, this probability with the factor Δt removed. Thus, the interspike interval distribution for a homogeneous Poisson spike train is an exponential. The most likely interspike intervals are short ones, and long intervals have a probability that falls exponentially as a function of their duration.

From the interspike interval distribution of a homogeneous Poisson spike train, we can compute the mean interspike interval,

$$\langle \tau \rangle = \int_0^\infty d\tau \, \tau r \exp(-r\tau) = \frac{1}{r} , \qquad (1.32)$$

and the variance of the interspike intervals,

$$\sigma_\tau^2 = \int_0^\infty d\tau \, \tau^2 r \exp(-r\tau) - \langle \tau \rangle^2 = \frac{1}{r^2} . \qquad (1.33)$$

The ratio of the standard deviation to the mean is called the coefficient of variation,

coefficient of variation C_V

$$C_V = \frac{\sigma_\tau}{\langle \tau \rangle} , \qquad (1.34)$$

and it takes the value 1 for a homogeneous Poisson process. This is a necessary, though not sufficient, condition to identify a Poisson spike train. Recall that the Fano factor for a Poisson process is also 1. For any renewal process, the Fano factor evaluated over long time intervals approaches the value C_V^2.

The Spike-Train Autocorrelation Function

The spike interval distribution measures the distribution of times between successive action potentials in a train. It is useful to generalize this concept and determine the distribution of times between any two spikes in

a train. This is called the spike-train autocorrelation function, and it is particularly useful for detecting patterns in spike trains, most notably oscillations. The spike-train autocorrelation function is the autocorrelation of the neural response function of equation 1.1 with its average over time and trials subtracted out. The time average of the neural response function, from equation 1.4, is the spike-count rate r, and the trial average of this quantity is $\langle r \rangle = \langle n \rangle / T$. Thus, the spike-train autocorrelation function is

*spike-train
autocorrelation
function $Q_{\rho\rho}$*

$$Q_{\rho\rho}(\tau) = \frac{1}{T} \int_0^T dt \, \langle (\rho(t) - \langle r \rangle) \, (\rho(t + \tau) - \langle r \rangle) \rangle \,. \qquad (1.35)$$

Because the average is subtracted from the neural response function in this expression, $Q_{\rho\rho}$ should really be called an autocovariance, not an autocorrelation, but in practice it isn't.

The spike-train autocorrelation function is constructed from data in the form of a histogram by dividing time into bins. The value of the histogram for a bin labeled with a positive or negative integer m is computed by determining the number of the times that any two spikes in the train are separated by a time interval lying between $(m - 1/2)\Delta t$ and $(m + 1/2)\Delta t$ with Δt the bin size. This includes all pairings, even between a spike and itself. We call this number N_m. If the intervals between the n^2 spike pairs in the train were uniformly distributed over the range from 0 to T, there would be $n^2 \Delta t / T$ intervals in each bin. This uniform term is removed from the autocorrelation histogram by subtracting $n^2 \Delta t / T$ from N_m for all m. The spike-train autocorrelation histogram is then defined by dividing the resulting numbers by T, so the value of the histogram in bin m is $H_m = N_m / T - n^2 \Delta t / T^2$. For small bin sizes, the $m = 0$ term in the histogram counts the average number of spikes, that is $N_m = \langle n \rangle$ and in the limit $\Delta t \to 0$, $H_0 = \langle n \rangle / T$ is the average firing rate $\langle r \rangle$. Because other bins have H_m of order Δt, the large $m = 0$ term is often removed from histogram plots. The spike-train autocorrelation function is defined as $H_m / \Delta t$ in the limit $\Delta t \to 0$, and it has the units of a firing rate squared. In this limit, the $m = 0$ bin becomes a δ function, $H_0 / \Delta t \to \langle r \rangle \delta(\tau)$.

As we have seen, the distribution of interspike intervals for adjacent spikes in a homogeneous Poisson spike train is exponential (equation 1.31). By contrast, the intervals between any two spikes (not necessarily adjacent) in such a train are uniformly distributed. As a result, the subtraction procedure outlined above gives $H_m = 0$ for all bins except for the $m = 0$ bin that contains the contribution of the zero intervals between spikes and themselves. The autocorrelation function for a Poisson spike train generated at a constant rate $\langle r \rangle = r$ is thus

$$Q_{\rho\rho}(\tau) = r\delta(\tau) \,. \qquad (1.36)$$

*cross-correlation
function*

A cross-correlation function between spike trains from two different neurons can be defined by analogy with the autocorrelation function by determining the distribution of intervals between pairs of spikes, one taken
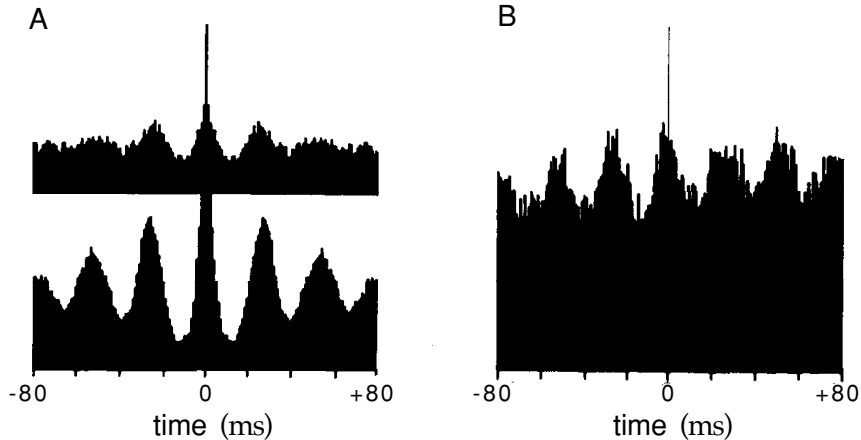
Figure 1.12 Autocorrelation and cross-correlation histograms for neurons in the primary visual cortex of a cat.   (A) Autocorrelation histograms for neurons recorded in the right (upper) and left (lower) hemispheres show a periodic pattern indicating oscillations at about 40 Hz. The lower diagram indicates stronger oscillations in the left hemisphere. (B) The cross-correlation histogram for these two neurons shows that their oscillations are synchronized with little time delay. (Adapted from Engel et al., 1991.)

from each train. The spike-train autocorrelation function is an even function of τ, $Q_{\rho\rho}(\tau) = Q_{\rho\rho}(-\tau)$, but the cross-correlation function is not necessarily even. A peak at zero interval in a cross-correlation function signifies that the two neurons are firing synchronously. Asymmetric shifts in this peak away from 0 result from fixed delays between the firing of the two neurons, and they indicate nonsynchronous but phase-locked firing. Periodic structure in either an autocorrelation or a cross-correlation function or histogram indicates that the firing probability oscillates. Such periodic structure is seen in the histograms of figure 1.12, showing 40 Hz oscillations in neurons of cat primary visual cortex that are roughly synchronized between the two cerebral hemispheres.

The Inhomogeneous Poisson Process

When the firing rate depends on time, different sequences of n spikes occur with different probabilities, and $p[t_1, t_2, \ldots, t_n]$ depends on the spike times. Because spikes are still generated independently by an inhomogeneous Poisson process, their times enter into $p[t_1, t_2, \ldots, t_n]$ only through the time-dependent firing rate $r(t)$. Assuming, as before, that the spike times are ordered $0 \le t_1 \le t_2 \le \ldots \le t_n \le T$, the probability density for n spike times (derived in appendix C) is

$$p[t_1, t_2, \ldots, t_n] = \exp\left(-\int_0^T dt\, r(t)\right) \prod_{i=1}^{n} r(t_i). \qquad (1.37)$$

This result applies if the spike times have been written in temporal order. If the spike times are not ordered, so that, for example, we are interested in the probability density for any spike occurring at the time t_1, not necessarily the first spike, this expression should be divided by a factor of $n!$ to account for the number of different possible orderings of spike times.

The Poisson Spike Generator

Spike sequences can be simulated by using some estimate of the firing rate, $r_{est}(t)$, predicted from knowledge of the stimulus, to drive a Poisson process. A simple procedure for generating spikes in a computer program is based on the fact that the estimated probability of firing a spike during a short interval of duration Δt is $r_{est}(t)\Delta t$. The program progresses through time in small steps of size Δt and generates, at each time step, a random number x_{rand} chosen uniformly in the range between 0 and 1. If $r_{est}(t)\Delta t > x_{rand}$ at that time step, a spike is fired; otherwise it is not.

For a constant firing rate, it is faster to compute spike times t_i for $i = 1, 2, \ldots n$ iteratively by generating interspike intervals from an exponential probability density (equation 1.31). If x_{rand} is uniformly distributed over the range between 0 and 1, the negative of its logarithm is exponentially distributed. Thus, we can generate spike times iteratively from the formula $t_{i+1} = t_i - \ln(x_{rand})/r$. Unlike the algorithm discussed in the previous paragraph, this method works only for constant firing rates. However, it can be extended to time-dependent rates by using a procedure called rejection sampling or spike thinning. The thinning technique requires a bound r_{max} on the estimated firing rate such that $r_{est}(t) \leq r_{max}$ at all times. We first generate a spike sequence corresponding to the constant rate r_{max} by iterating the rule $t_{i+1} = t_i - \ln(x_{rand})/r_{max}$. The spikes are then thinned by generating another x_{rand} for each i and removing the spike at time t_i from the train if $r_{est}(t_i)/r_{max} < x_{rand}$. If $r_{est}(t_i)/r_{max} \geq x_{rand}$, spike i is retained. Thinning corrects for the difference between the estimated time-dependent rate and the maximum rate.

Figure 1.13 shows an example of a model of an orientation-selective V1 neuron constructed in this way. In this model, the estimated firing rate is determined from the response tuning curve of figure 1.5B,

$$r_{est}(t) = f(s(t)) = r_{max} \exp\left(-\frac{1}{2}\left(\frac{s(t) - s_{max}}{\sigma_f}\right)^2\right). \qquad (1.38)$$

This is an extremely simplified model of response dynamics, because the firing rate at any given time depends only on the value of the stimulus at that instant of time and not on its recent history. Models that allow for a dependence of firing rate on stimulus history are discussed in chapter 2. In figure 1.13, the orientation angle increases in a sequence of steps. The firing rate follows these changes, and the Poisson process generates an irregular firing pattern that reflects the underlying rate but varies from trial to trial.
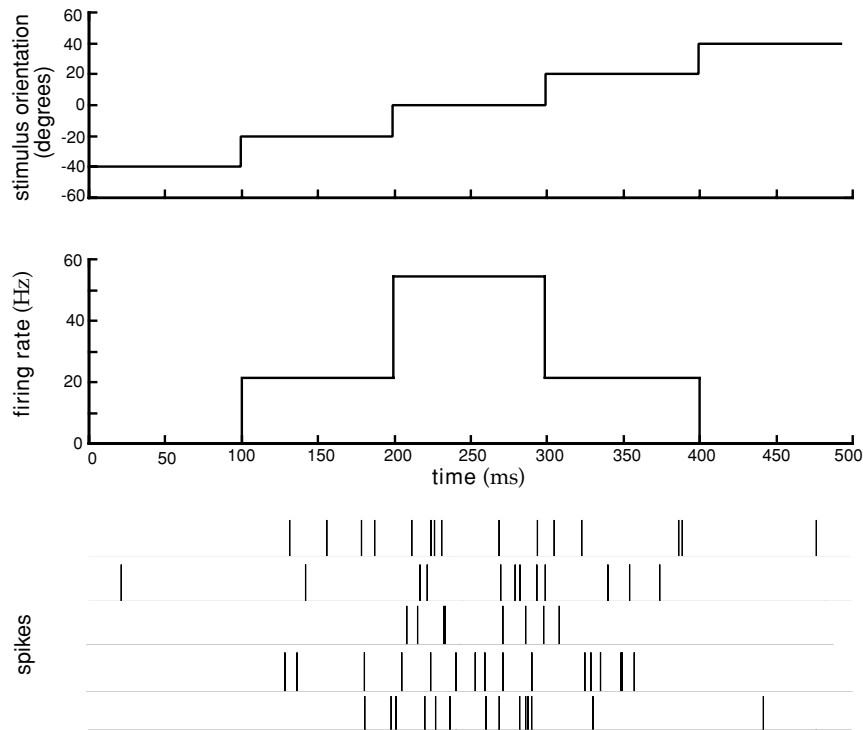
Figure 1.13 Model of an orientation-selective neuron. The orientation angle (top panel) was increased from an initial value of -40° by 20° every 100 ms. The firing rate (middle panel) was used to generate spikes (bottom panel) using a Poisson spike generator. The bottom panel shows spike sequences generated on five different trials.

Certain features of neuronal firing violate the independence assumption that forms the basis of the Poisson model, at least if a constant firing rate is used. We have already mentioned the absolute and relative refractory periods, which are periods of time following the generation of an action potential when the probability of a spike occurring is greatly or somewhat reduced. Frequently, these are most prominent features of real neuronal spike trains that are not captured by a Poisson model. Refractory effects can be incorporated into a Poisson model of spike generation by setting the firing rate to 0 immediately after a spike is fired, and then letting it return to its predicted value according to some dynamic rule such as an exponential recovery.

Comparison with Data

The Poisson process is simple and useful, but does it match data on neural response variability? To address this question, we examine Fano factors, interspike interval distributions, and coefficients of variation.
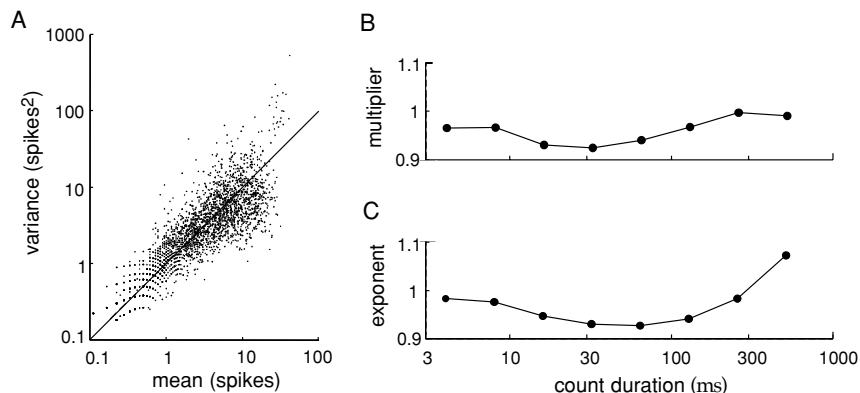
Figure 1.14 Variability of MT neurons in alert macaque monkeys responding to moving visual images. (A) Variance of the spike counts for a 256 ms counting period plotted against the mean spike count. The straight line is the prediction of the Poisson model. Data are from 94 cells recorded under a variety of stimulus conditions. (B) The multiplier A in the relationship between spike-count variance and mean as a function of the duration of the counting interval. (C) The exponent B in this relation as a function of the duration of the counting interval. (Adapted from O'Keefe et al., 1997.)

The Fano factor describes the relationship between the mean spike count over a given interval and the spike-count variance. Mean spike counts $\langle n \rangle$ and variances σ_n^2 from a wide variety of neuronal recordings have been fitted to the equation $\sigma_n^2 = A\langle n \rangle^B$, and the multiplier A and exponent B have been determined. The values of both A and B typically lie between 1.0 and 1.5. Because the Poisson model predicts $A = B = 1$, this indicates that the data show a higher degree of variability than the Poisson model would predict. However, many of these experiments involve anesthetized animals, and it is known that response variability is higher in anesthetized than in alert animals.

area MT

Figure 1.14 shows data for spike-count means and variances extracted from recordings of MT neurons in alert macaque monkeys using a number of different stimuli. The MT (medial temporal) area is a visual region of the primate cortex where many neurons are sensitive to image motion. The individual means and variances are scattered in figure 1.14A, but they cluster around the diagonal which is the Poisson prediction. Similarly, the results show A and B values close to 1, the Poisson values (figure 1.14B). Of course, many neural responses cannot be described by Poisson statistics, but it is reassuring to see a case where the Poisson model seems a reasonable approximation. As mentioned previously, when spike trains are not described very accurately by a Poisson model, refractory effects are often the primary reason.

Interspike interval distributions are extracted from data as interspike interval histograms by counting the number of intervals falling in discrete time bins. Figure 1.15A presents an example from the responses of a non-bursting cell in area MT of a monkey in response to images consisting of
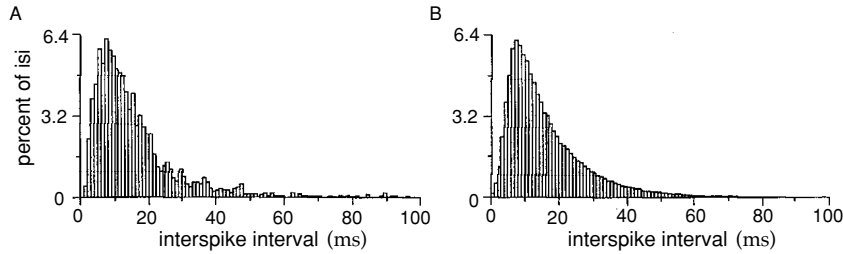
Figure 1.15 (A) Interspike interval distribution from an MT neuron responding to a moving, random-dot image. The probability of interspike intervals falling into the different bins, expressed as a percentage, is plotted against interspike interval. (B) Interspike interval histogram generated from a Poisson model with a stochastic refractory period. (Adapted from Bair et al., 1994.)

randomly moving dots with a variable amount of coherence imposed on their motion (see chapter 3 for a more detailed description). For interspike intervals longer than about 10 ms, the shape of this histogram is exponential, in agreement with equation 1.31. However, for shorter intervals there is a discrepancy. While the homogeneous Poisson distribution of equation 1.31 rises for short interspike intervals, the experimental results show a rapid decrease. This is the result of refractoriness making short interspike intervals less likely than the Poisson model would predict. Data on interspike intervals can be fitted more accurately by a gamma distribution, *gamma distribution*

$$p[\tau] = \frac{r(r\tau)^k \exp(-r\tau)}{k!} \tag{1.39}$$

with $k > 0$, than by the exponential distribution of the Poisson model, which has $k = 0$.

Figure 1.15B shows a theoretical histogram obtained by adding a refractory period of variable duration to the Poisson model. Spiking was prohibited during the refractory period, and then was described once again by a homogeneous Poisson process. The refractory period was randomly chosen from a Gaussian distribution with a mean of 5 ms and a standard deviation of 2 ms (only random draws that generated positive refractory periods were included). The resulting interspike interval distribution of figure 1.15B agrees quite well with the data.

C_V values extracted from the spike trains of neurons recorded in monkeys from area MT and primary visual cortex (V1) are shown in figure 1.16. The data have been divided into groups based on the mean interspike interval, and the coefficient of variation is plotted as a function of this mean interval, equivalent to $1/\langle r \rangle$. Except for short mean interspike intervals, the values are near 1, although they tend to cluster slightly lower than 1, the Poisson value. The small C_V values for short interspike intervals are due to the refractory period. The solid curve is the prediction of a Poisson model with refractoriness.

The Poisson model with refractoriness provides a reasonably good description of a significant amount of data, especially considering its sim-
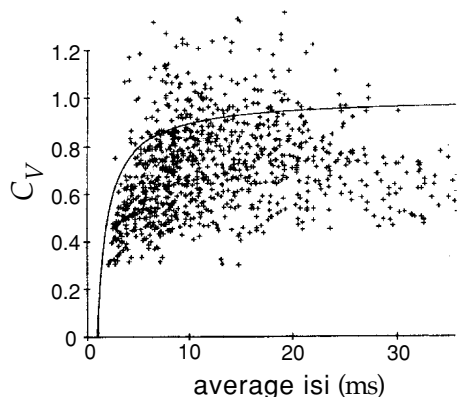
Figure 1.16 Coefficients of variation for a large number of V1 and MT neurons plotted as a function of mean interspike interval. The solid curve is the result of a Poisson model with a refractory period. (Adapted from Softky and Koch, 1992.)

plicity. However, there are cases in which the accuracy in the timing and numbers of spikes fired by a neuron is considerably higher than would be implied by Poisson statistics. Furthermore, even when it successfully describes data, the Poisson model does not provide a mechanistic explanation of neuronal response variability. Spike generation, by itself, is highly reliable in real neurons. Figure 1.17 compares the response of V1 cells to constant current injection in vivo and in vitro. The in vitro response is a regular and reproducible spike train (left panel). The same current injection paradigm applied in vivo produces a highly irregular pattern of firing (center panel) similar to the response to a moving bar stimulus (right panel). Although some of the basic statistical properties of firing variability may be captured by the Poisson model of spike generation, the spike-generating mechanism itself in real neurons is clearly not responsible for the variability. We explore ideas about possible sources of spike-train variability in chapter 5.

Some neurons fire action potentials in clusters or bursts of spikes that cannot be described by a Poisson process with a fixed rate. Bursting can be included in a Poisson model by allowing the firing rate to fluctuate in order to describe the high rate of firing during a burst. Sometimes the distribution of bursts themselves can be described by a Poisson process (such a doubly stochastic process is called a Cox process).

1.5   The Neural Code

The nature of the neural code is a topic of intense debate within the neuroscience community. Much of the discussion has focused on whether neurons use rate coding or temporal coding, often without a clear definition of what these terms mean. We feel that the central issue in neural coding is whether individual action potentials and individual neurons encode inde-
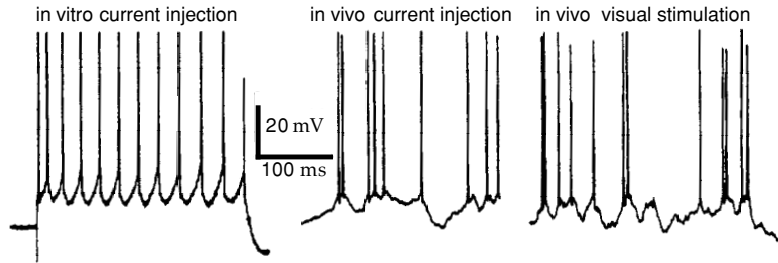
in vitro current injection   in vivo current injection   in vivo visual stimulation

20 mV

100 ms

Figure 1.17 Intracellular recordings from cat V1 neurons. The left panel is the response of a neuron in an in vitro slice preparation to constant current injection. The center and right panels show recordings from neurons in vivo responding to either injected current (center) or a moving visual image (right). (Adapted from Holt et al., 1996.)

pendently of each other, or whether correlations between different spikes and different neurons carry significant amounts of information. We therefore contrast independent-spike and independent-neuron codes with correlation codes before addressing the issue of temporal coding.

Independent-Spike, Independent-Neuron, and Correlation Codes

The neural response, and its relation to the stimulus, are completely characterized by the probability distribution of spike times as a function of the stimulus. If spike generation can be described as an inhomogeneous Poisson process, this probability distribution can be computed from the time-dependent firing rate $r(t)$, using equation 1.37. In this case, $r(t)$ contains all the information about the stimulus that can be extracted from the spike train, and the neural code could reasonably be called a rate code. Unfortunately, this definition does not agree with common usage. Instead, we will call a code based solely on the time-dependent firing rate an independent-spike code. This refers to the fact that the generation of each spike is independent of all the other spikes in the train. If individual spikes do not encode independently of each other, we call the code a correlation code, because correlations between spike times may carry additional information. In reality, information is likely to be carried both by individual spikes and through correlations, and some arbitrary dividing line must be established to characterize the code. Identifying a correlation code should require that a significant amount of information be carried by correlations, for example, as much as is carried by the individual spikes.

independent-spike code

correlation code

A simple example of a correlation code would occur if significant amounts of information about a stimulus were carried by interspike intervals. In this case, if we considered spike times individually, independently of each other, we would miss the information carried by the intervals between them. This is just one example of a correlation code. Information could be carried by more complex relationships between spikes.

Independent-spike codes are much simpler to analyze than correlation codes, and most work on neural coding assumes spike independence. When careful studies have been done, it has been found that some information is carried by correlations between two or more spikes, but this information is rarely larger than 10% of the information carried by spikes considered independently. Of course, it is possible that, due to our ignorance of the "real" neural code, we have not yet uncovered or examined the types of correlations that are most significant for neural coding. Although this is not impossible, we view it as unlikely and feel that the evidence for independent-spike coding, at least as a fairly accurate approximation, is quite convincing.

The discussion to this point has focused on information carried by single neurons, but information is typically encoded by neuronal populations. When we study population coding, we must consider whether individual neurons act independently, or whether correlations between different neurons carry additional information. The analysis of population coding is easiest if the response of each neuron is considered statistically independent, and such independent-neuron coding is typically assumed in the analysis of population codes (chapter 3). The independent-neuron hypothesis does not mean that the spike trains of different neurons are not combined into an ensemble code. Rather, it means that they can be combined without taking correlations into account. To test the validity of this assumption, we must ask whether correlations between the spiking of different neurons provide additional information about a stimulus that cannot be obtained by considering all of their firing patterns individually.

independent-neuron code

synchrony and oscillations

Synchronous firing of two or more neurons is one mechanism for conveying information in a population correlation code. Rhythmic oscillations of population activity provide another possible mechanism, as discussed below. Both synchronous firing and oscillations are common features of the activity of neuronal populations. However, the existence of these features is not sufficient for establishing a correlation code, because it is essential to show that a significant amount of information is carried by the resulting correlations. The assumption of independent-neuron coding is a useful simplification that is not in gross contradiction with experimental data, but it is less well established and more likely to be challenged in the future than the independent-spike hypothesis.

hippocampal place cells

Place-cell coding of spatial location in the rat hippocampus is an example in which at least some additional information appears to be carried by correlations between the firing patterns of neurons in a population. The hippocampus is a structure located deep inside the temporal lobe that plays an important role in memory formation and is involved in a variety of spatial tasks. The firing rates of many hippocampal neurons, recorded when a rat is moving around a familiar environment, depend on the location of the animal and are restricted to spatially localized areas called the place fields of the cells. In addition, when a rat explores an environment, hippocampal neurons fire collectively in a rhythmic pattern with a frequency in the theta range, 7-12 Hz. The spiking time of an individual
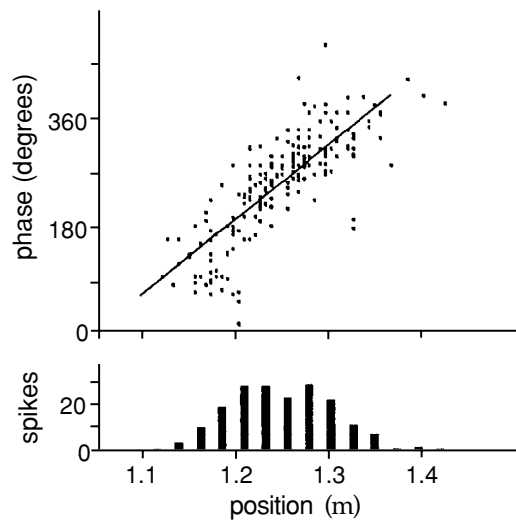
Figure 1.18 Position versus phase for a hippocampal place cell. Each dot in the upper figure shows the phase of the theta rhythm plotted against the position of the animal at the time when a spike was fired. The linear relation shows that information about position is contained in the relative phase of firing. The lower plot is a conventional place field tuning curve of spike count versus position. (Adapted from O'Keefe and Recce, 1993.)

place cell relative to the phase of the population theta rhythm gives additional information about the location of the rat not provided by place cells considered individually. The relationship between location and phase of place-cell firing shown in figure 1.18 means, for example, that we can distinguish two locations on opposite sides of the peak of a single neuron's tuning curve that correspond to the same firing rate, by knowing when the spikes occurred relative to the theta rhythm. However, the amount of additional information carried by correlations between place-field firing and the theta rhythm has not been fully quantified.

Temporal Codes

The concept of temporal coding arises when we consider how precisely we must measure spike times to extract most of the information from a neuronal response. This precision determines the temporal resolution of the neural code. A number of studies have found that this temporal resolution is on a millisecond time scale, indicating that precise spike timing is a significant element in neural encoding. Similarly, we can ask whether high-frequency firing-rate fluctuations carry significant information about a stimulus. When precise spike timing or high-frequency firing-rate fluctuations are found to carry information, the neural code is often identified as a temporal code.

The temporal structure of a spike train or firing rate evoked by a stimulus is determined both by the dynamics of the stimulus and by the nature of the neural encoding process. Stimuli that change rapidly tend to generate precisely timed spikes and rapidly changing firing rates no matter what neural coding strategy is being used. Temporal coding refers to (or should refer to) temporal precision in the response that does not arise solely from the dynamics of the stimulus, but that nevertheless relates to properties of the stimulus. The interplay between stimulus and encoding dynamics makes the identification of a temporal code difficult.

The issue of temporal coding is distinct and independent from the issue of independent-spike coding discussed above. If the independent-spike hypothesis is valid, the temporal character of the neural code is determined by the behavior of $r(t)$. If $r(t)$ varies slowly with time, the code is typically called a rate code, and if it varies rapidly, the code is called temporal. Figure 1.19 provides an example of different firing-rate behaviors for a neuron in area MT of a monkey recorded over multiple trials with three different stimuli (consisting of moving random dots). The activity in the top panel would typically be regarded as reflecting rate coding, and the activity in the bottom panel as reflecting temporal coding. However, the identification of rate and temporal coding in this way is ambiguous because it is not obvious what criterion should be used to characterize the changes in $r(t)$ as slow or rapid.

One possibility is to use the spikes to distinguish slow from rapid, so that a temporal code is identified when peaks in the firing rate occur with roughly the same frequency as the spikes themselves. In this case, each peak corresponds to the firing of only one, or at most a few action potentials. While this definition makes intuitive sense, it is problematic to extend it to the case of population coding. When many neurons are involved, any single neuron may fire only a few spikes before its firing rate changes, but collectively the population may produce a large number of spikes over the same time period. Thus, by this definition, a neuron that appears to employ a temporal code may be part of a population that does not.

Another proposal is to use the stimulus, rather than the response, to establish what makes a temporal code. In this case, a temporal code is defined as one in which information is carried by details of spike timing on a scale shorter than the fastest time characterizing variations of the stimulus. This requires that information about the stimulus be carried by Fourier components of $r(t)$ at frequencies higher than those present in the stimulus. Many of the cases where a temporal code has been reported using spikes to define the nature of the code would be called rate codes if the stimulus were used instead.

The debate between rate and temporal coding dominates discussions about the nature of the neural code. Determining the temporal resolution of the neural code is clearly important, but much of this debate seems uninformative. We feel that the central challenge is to identify relationships
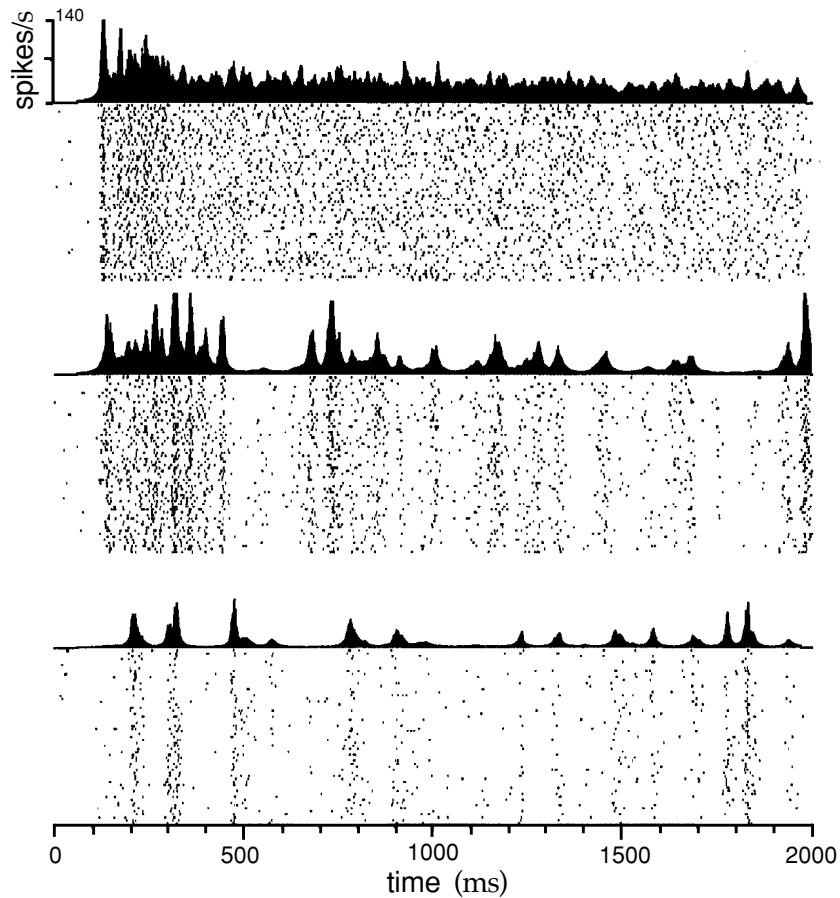
Figure 1.19 Time-dependent firing rates for different stimulus parameters. The rasters show multiple trials during which an MT neuron responded to the same moving, random-dot stimulus. Firing rates, shown above the raster plots, were constructed from the multiple trials by counting spikes within discrete time bins and averaging over trials. The three different results are from the same neuron but using different stimuli. The stimuli were always patterns of moving random dots, but the coherence of the motion was varied (see chapter 3 for more information about this stimulus). (Adapted from Bair and Koch, 1996.)

between the firing patterns of different neurons in a responding population and to understand their significance for neural coding.

1.6   Chapter Summary

With this chapter, we have begun our study of the way that neurons encode information using spikes. We used a sequence of δ functions, the neural response function, to represent a spike train and defined three types of firing rates: the time-dependent firing rate $r(t)$, the spike-count rate r,

and the average firing rate $\langle r \rangle$. In the discussion of how the firing rate $r(t)$ could be extracted from data, we introduced the important concepts of a linear filter and a kernel acting as a sliding window function. The average firing rate expressed as a function of a static stimulus parameter is called the response tuning curve, and we presented examples of Gaussian, cosine, and sigmoidal tuning curves. Spike-triggered averages of stimuli, or reverse correlation functions, were introduced to characterize the selectivity of neurons to dynamic stimuli. The homogeneous and inhomogeneous Poisson processes were presented as models of stochastic spike sequences. We defined correlation functions, auto- and cross-correlations, and power spectra, and used the Fano factor, interspike-interval histogram, and coefficient of variation to characterize the stochastic properties of spiking. We concluded with a discussion of independent-spike and independent-neuron codes versus correlation codes, and of the temporal precision of spike timing as addressed in discussions of temporal coding.

1.7   Appendices

A: The Power Spectrum of White Noise

The Fourier transform of the stimulus autocorrelation function (see the Mathematical Appendix),

$$\tilde{Q}_{ss}(\omega) = \frac{1}{T} \int_{-T/2}^{T/2} d\tau \, Q_{ss}(\tau) \exp(i\omega\tau) \,, \tag{1.40}$$

power spectrum   is called the power spectrum. Because we have defined the stimulus as periodic outside the range of the trial T, we have used a finite-time Fourier transform and ω should be restricted to values that are integer multiples of $2\pi/T$. We can compute the power spectrum for a white-noise stimulus using the fact that $Q_{ss}(\tau) = \sigma_s^2 \delta(\tau)$ for white noise,

$$\tilde{Q}_{ss}(\omega) = \frac{\sigma_s^2}{T} \int_{-T/2}^{T/2} d\tau \, \delta(\tau) \exp(i\omega\tau) = \frac{\sigma_s^2}{T} \,. \tag{1.41}$$

This is the defining characteristic of white noise; its power spectrum is independent of frequency.

Using the definition of the stimulus autocorrelation function, we can also write

$$
\begin{aligned}
\tilde{Q}_{ss}(\omega) &= \frac{1}{T} \int_0^T dt \, s(t) \frac{1}{T} \int_{-T/2}^{T/2} d\tau \, s(t+\tau) \exp(i\omega\tau) \\
&= \frac{1}{T} \int_0^T dt \, s(t) \exp(-i\omega t) \frac{1}{T} \int_{-T/2}^{T/2} d\tau \, s(t+\tau) \exp(i\omega(t+\tau)) \,.
\end{aligned}
\tag{1.42}
$$

The first integral on the right side of the second equality is the complex conjugate of the Fourier transform of the stimulus,

$$\tilde{s}(\omega) = \frac{1}{T} \int_0^T d\tau\, s(t) \exp(i\omega t)\,. \tag{1.43}$$

The second integral, because of the periodicity of the integrand (when ω is an integer multiple of $2\pi/T$) is equal to $\tilde{s}(\omega)$. Therefore,

$$\tilde{Q}_{ss}(\omega) = |\tilde{s}(\omega)|^2\,, \tag{1.44}$$

which provides another definition of the stimulus power spectrum. It is the absolute square of the Fourier transform of the stimulus.

Although equations 1.40 and 1.44 are both sound, they do not provide a statistically efficient method of estimating the power spectrum of discrete approximations to white-noise sequences generated by the methods described in this chapter. That is, the apparently natural procedure of taking a white-noise sequence $s(m\Delta t)$ for $m = 1, 2, \ldots, T/\Delta t$, and computing the square amplitude of its Fourier transform at frequency ω,

$$\frac{\Delta T}{T} \left| \sum_{m=1}^{T/\Delta t} s(t) \exp(-i\omega m\Delta t) \right|^2\,,$$

is a biased and extremely noisy way of estimating $\tilde{Q}_{ss}(\omega)$. This estimator is called the periodogram. The statistical problems with the periodogram, *periodogram* and some of the many suggested solutions, are discussed in almost any textbook on spectral analysis (see, e.g., Percival and Waldron, 1993).

B: Moments of the Poisson Distribution

The average number of spikes generated by a Poisson process with constant rate r over a time T is

$$\langle n \rangle = \sum_{n=0}^{\infty} n P_T[n] = \sum_{n=0}^{\infty} \frac{n(rT)^n}{n!} \exp(-rT)\,, \tag{1.45}$$

and the variance in the spike count is

$$\sigma_n^2(T) = \sum_{n=0}^{\infty} n^2 P_T[n] - \langle n \rangle^2 = \sum_{n=0}^{\infty} \frac{n^2(rT)^n}{n!} \exp(-rT) - \langle n \rangle^2\,. \tag{1.46}$$

To compute these quantities, we need to calculate the two sums appearing in these equations. A good way to do this is to compute the moment-generating function *moment-generating function*

$$g(\alpha) = \sum_{n=0}^{\infty} \frac{(rT)^n \exp(\alpha n)}{n!} \exp(-rT)\,. \tag{1.47}$$

The kth derivative of g with respect to α, evaluated at the point $\alpha = 0$, is

$$\frac{d^k g}{d\alpha^k}\bigg|_{\alpha=0} = \sum_{n=0}^{\infty} \frac{n^k (rT)^n}{n!} \exp(-rT) , \qquad (1.48)$$

so once we have computed g, we need to calculate only its first and second derivatives to determine the sums we need. Rearranging the terms a bit, and recalling that $\exp(z) = \sum z^n / n!$, we find

$$g(\alpha) = \exp(-rT) \sum_{n=0}^{\infty} \frac{\left(rT \exp(\alpha)\right)^n}{n!} = \exp(-rT) \exp\left(rTe^{\alpha}\right) . \qquad (1.49)$$

The derivatives are then

$$\frac{dg}{d\alpha} = rTe^{\alpha} \exp(-rT) \exp(rTe^{\alpha}) \qquad (1.50)$$

and

$$\frac{d^2 g}{d\alpha^2} = (rTe^{\alpha})^2 \exp(-rT) \exp(rTe^{\alpha}) + rTe^{\alpha} \exp(-rT) \exp(rTe^{\alpha}) . \qquad (1.51)$$

Evaluating these at $\alpha = 0$ and putting the results into equations 1.45 and 1.46 gives the results $\langle n \rangle = rT$ and $\sigma_n^2(T) = (rT)^2 + rT - (rT)^2 = rT$.

C: Inhomogeneous Poisson Statistics

The probability density for a particular spike sequence with spike times t_i for $i = 1, 2, \ldots, n$ is obtained from the corresponding probability distribution by multiplying the probability that the spikes occur when they do by the probability that no other spikes occur. We begin by computing the probability that no spikes are generated during the time interval from t_i to t_{i+1} between two adjacent spikes. We determine this by dividing the interval into M bins of size Δt and setting $M\Delta t = t_{i+1} - t_i$. We will ultimately take the limit $\Delta t \to 0$. The firing rate during bin m within this interval is $r(t_i + m\Delta t)$. Because the probability of firing a spike in this bin is $r(t_i + m\Delta t)\Delta t$, the probability of not firing a spike is $1 - r(t_i + m\Delta t)\Delta t$. To have no spikes during the entire interval, we must string together M such bins, and the probability of this occurring is the product of the individual probabilities,

$$P[\text{no spikes}] = \prod_{m=1}^{M} (1 - r(t_i + m\Delta t)\Delta t) . \qquad (1.52)$$

We evaluate this expression by taking its logarithm,

$$\ln P[\text{no spikes}] = \sum_{m=1}^{M} \ln (1 - r(t_i + m\Delta t)\Delta t) , \qquad (1.53)$$

using the fact that the logarithm of a product is the sum of the logarithms of the multiplied terms. Using the approximation $\ln(1 - r(t_i + m\Delta t)\Delta t) \approx -r(t_i + m\Delta t)\Delta t$, valid for small Δt, we can simplify this to

$$\ln P[\text{no spikes}] = -\sum_{m=1}^{M} r(t_i + m\Delta t)\Delta t. \tag{1.54}$$

In the limit $\Delta t \to 0$, the approximation becomes exact and this sum becomes the integral of $r(t)$ from t_i to t_{i+1},

$$\ln P[\text{no spikes}] = -\int_{t_i}^{t_{i+1}} dt\, r(t). \tag{1.55}$$

Exponentiating this equation gives the result we need,

$$P[\text{no spikes}] = \exp\left(-\int_{t_i}^{t_{i+1}} dt\, r(t)\right). \tag{1.56}$$

The probability density $p[t_1, t_2, \ldots, t_n]$ is the product of the densities for the individual spikes and the probabilities of not generating spikes during the interspike intervals, between time 0 and the first spike, and between the time of the last spike and the end of the trial period:

$$p[t_1, t_2, \ldots, t_n] = \exp\left(-\int_0^{t_1} dt\, r(t)\right)\exp\left(-\int_{t_n}^{T} dt\, r(t)\right) \times$$
$$r(t_n)\prod_{i=1}^{n-1} r(t_i)\exp\left(-\int_{t_i}^{t_{i+1}} dt\, r(t)\right). \tag{1.57}$$

The exponentials in this expression all combine because the product of exponentials is the exponential of the sum, so the different integrals in this sum add up to form a single integral:

$$\exp\left(-\int_0^{t_1} dt\, r(t)\right)\exp\left(-\int_{t_n}^{T} dt\, r(t)\right)\prod_{i=1}^{n-1}\exp\left(-\int_{t_i}^{t_{i+1}} dt\, r(t)\right)$$
$$= \exp\left(-\left(\int_0^{t_1} dt\, r(t) + \sum_{i=1}^{n-1}\int_{t_i}^{t_{i+1}} dt\, r(t) + \int_{t_n}^{T} dt\, r(t)\right)\right)$$
$$= \exp\left(-\int_0^{T} dt\, r(t)\right). \tag{1.58}$$

Substituting this into 1.57 gives the result in equation 1.37.

1.8 Annotated Bibliography

Braitenberg & Schuz (1991) provides some of the quantitative measures of neuroanatomical properties of cortex that we quote. **Rieke et al. (1997)**

describes the analysis of spikes and the relationships between neural responses and stimuli, and is a general reference for material we present in chapters 1–4. **Gabbiani & Koch (1998)** provides another account of some of this material. The mathematics underlying point processes, the natural statistical model for spike sequences, is found in **Cox (1962)** and **Cox & Isham (1980)**, including the relationship between the Fano factor and the coefficient of variation. A general analysis of histogram representations appears in **Scott (1992)**, and white-noise and filtering techniques (our analysis of which continues in chapter 2) are described in de Boer & Kuyper (1968), **Marmarelis & Marmarelis (1978)**, and **Wiener (1958)**. Berry & Meister (1998) discuss the effects of refractoriness on patterns of spiking.

In chapters 1 and 3, we discuss two systems associated with studies of spike encoding; the H1 neuron in the visual system of flies, reviewed by **Rieke et al. (1997)**, and area MT of monkeys, discussed by Parker & Newsome (1998). **Wandell (1995)** introduces orientation and disparity tuning, relevant to examples presented in this chapter.