CHAPTER 5

Signal Detection

Nothing is great or little otherwise than by comparison.
—Jonathan Swift, *Gulliver's Travels*, 1726

Given the transmission of a stimulus to an organism, the next problem is that of signal detection. Is the intensity of the stimulus sufficient? Does noise interfere with detection? What features of the stimulus are detected? What strategies are useful to the receiving organism? A large body of theory about signal detection has been developed, primarily for efficient design of electronic communications. This chapter presents some basic ideas useful to sensory ecology. Here the term *signal* will be used in the general sense of a meaningful stimulus from any source.

Signal detection can be said to occur when the output of a sensory system (the **subjective intensity**) exceeds some threshold value. The **objective intensity** impinging on the receiver is determined by the intensity of the signal transmitted by the sender and the degree to which the signal is attenuated in the environment between the sender and receiver. The subjective intensity, which controls behavior, is determined by the objective intensity, the sensitivity of the receptor, and the signal-processing systems in the receiving organism. The threshold is usually set by the receiver at an appropriate level in regard to the level of noise in the system.

5-1 INTENSITY NOISE

There are many sources of noise that affect subjective intensity and in various situations different noise sources dominate. In general, it is useful to distinguish three types of noise: receptor, channel, and environmental. **Receptor noise** is present in all receptor systems as a result of the thermal motion of molecules. Man can reduce the noise in fabricated infrared detectors by cooling them, but no other organism is known to employ such a strategy. Instead, receptor noise is reduced by averaging over a sufficiently large number of molecules or other elements. Examples of receptor noise are the hiss in radio output when one is tuned between stations and the dark current of a photomultiplier tube. In humans we might imagine that visual phantoms and hallucinations arise from noise in receptors or signal-processing parts of the nervous system.

The term **channel noise** is used to describe noise caused by fluctuations in the signal. For example, photon noise results from statistical fluctuations in the number of photons impinging on a detector. Our vision in dim light is limited by this type of noise. At a distance, sound intensity is modulated randomly by fluctuations in density and velocity of the intervening medium, whether it is air or water (see Section 9-4). The ability of bacteria to detect chemical gradients is limited by random fluctuations in the rate at which molecules diffuse into and out of their vicinity (see Sections 7-3 and 17-2). In many such cases channel noise, as well as receptor noise, is nearly random and can be treated by statistical analysis.

In contrast, **environmental noise** is created by competing sources, whose output is usually not random. Some examples are the difficulty of detecting the song of a particular bird in the presence of other bird songs, or a flying moth's mistaking a streetlight for the moon, which it uses as a directional reference. Environmental noise usually limits the detectability of temperature and sound stimuli (see Chapters 6 and 9). In most cases environmental noise is usually specific to a given situation, so that no general analysis can be made.

5-2 DETECTION THEORY

Theories of signal detection have been extensively developed for use in electronic communications and radar (Selin 1965; Poor 1988). It may be noted that detection theory is closely related to statistical hypothesis testing, which is more familiar to biologists.

In general, a channel has noise even in the absence of a signal. As depicted in Figure 5-1, any of a range of intensities may be observed, with varying probability. In many cases the probability distribution is likely to be approximately normal (Gaussian). When a signal is present, the distribution is shifted and may or may not retain the same shape. The problem is to decide whether or not a signal is indeed present when a given intensity is observed. In simple cases the problem reduces to one of determining a threshold intensity upon which to base the decision. If the observed intensity is below the threshold, one concludes that a signal is not present. Higher intensities are interpreted as meaning a signal is present. If there is no overlap in the two probability distributions, an intermediate threshold intensity can be chosen as a threshold which produces completely reliable decisions and noise poses no problem. The degree of overlap is usually measured by the **signal-to-noise** (S/N) ratio. In Figure 5-1, the signal (S) is a measure of the distance between the mean intensities with and without the signal, and the noise (N) is a measure of the width of the probability distributions.

Detection theory addresses the question of the best strategy to adopt when the two distributions do overlap. If the probability distributions in Figure 5-1 are known, the obvious strategy is to set the threshold, I_{Th}, at the intensity level where the two distributions intersect. By doing so, for any

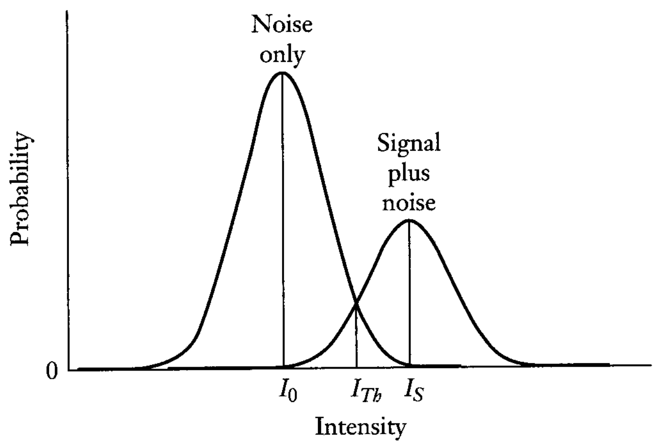

Figure 5-1 The probability of observing a particular intensity in the presence and absence of a signal. The areas under the two curves must sum to 1, since the signal is either present or not. The difference between I_S and I_0 is the strength of the signal, and the width of the distributions is a measure of the noise. The intensity at I_{Th}, where the curves intersect, is the optimal threshold. Intensities above I_{Th} are more likely to occur when a signal is present, intensities below it more likely when one is absent.

observed intensity the more probable case of the presence or absence of the signal is chosen, which maximizes the overall probability of reaching correct decisions. If the distributions overlap, it is not possible to be completely reliable. Reliability can be improved to any desired level by making additional observations, if possible. This strategy is analogous to that of increasing the integration time of a receptor, as described shortly. A more sophisticated strategy would be to repeat the observations only when ambiguous, intermediate intensities are observed. For example, humans often sniff several times after detecting a hint of an important smell.

Usually, a situation is not so well-defined that an optimal threshold intensity can be chosen this easily. In particular, the probability distributions may not be known. Even if there is knowledge about the noise in the channel, it is also necessary to know the prior probability that a signal is present, which would influence the relative height of the two probability distributions and hence their point of intersection. Another complication is that the cost of missing a signal may be different from the cost of falsely concluding that it is present. For example, the cost of missing an alarm call may be death, whereas the cost of falsely sensing an alarm is only wasted time and energy. In such cases it is desirable to minimize the cost rather than the number of errors.

The simplest quantitative way to incorporate all these considerations is to consider a channel that has a mean intensity of I_0 in the absence of a signal and I_S in its presence, and has noise that has a normal distribution with standard deviation I_{SD}, whether or not a signal is present. Furthermore, assume that the prior probability of a signal's being present is P_S and that the cost of an error minus any cost of a correct determination is c_0 when no signal is present and c_S when one is present. Then the threshold that minimizes overall cost is calculated as (Poor 1988, 17)

$$I_{Th} = \frac{I_0 + I_S}{2} + \frac{I_{SD}^2}{I_S - I_0} \ln\left(\frac{(1 - P_S)\, c_0}{P_S\, c_S}\right) \tag{5-1}$$

Note that when the probabilities and costs are equal for the two situations $P_S = 0.5$ and $c_0 = c_S$, the second term goes to zero and, as anticipated, the threshold becomes the midpoint, $(I_0 + I_S)/2$, between the two distributions, which, by assumption, have an equal width. In the absence of knowledge about the prior probability of a signal's being present, the threshold that minimizes the cost with the worst possible prior probability is also the midpoint (Poor 1988, 28).

The trade-off between the two types of errors is shown in Figure 5-2. For a given signal-to-noise ratio, the receiver can choose a threshold that will give any combination of "false-alarm" and "miss" errors that falls on the appropriate curve. Both types of error can be reduced only by increasing the signal-to-noise ratio. In the absence of signal ($S/N = 0$) the only strategy available is pure guessing, but the receiver could guess the signal to be always present, always absent, half and half, or some other combination. All these possibilities lie on the diagonal in Figure 5-2. If a clear signal ($S/N >> 1$) is available, correct decisions can always be made, and the performance will correspond to that in the upper-left-hand corner. For ambiguous signals ($S/N \approx 1$), performance will be confined to the appropriate curve in the upper left half of the figure, but the receiver will be able to operate at any point along that curve.

For a more specific example, consider a channel in which events occur randomly in time. Such events might be the arrival of photons or molecules. If I is the average rate at which such events are received over a time interval t, the number of events is $N_t = I t$. Then the root-mean-square (rms) deviation, or the standard deviation, in N_t is $N_t^{1/2}$, and the relative rms deviation is $N_t^{1/2}/N_t = N_t^{-1/2}$. Consequently, the relative intensity noise is expressed as

$$\frac{I_{SD}}{I} = (I\,t)^{-1/2} \tag{5-2}$$

and the spread in the distribution of intensity can be reduced to an arbitrarily small value by increasing either the integration time or the stimulus flux to increase the number of events observed. Note, however, that the noise is reduced only in proportion to the square root of the stimulus intensity or sample time. Increasing sample size is the basic reason our visual processes are relatively slow and is why nocturnal animals have large eyes (see Sections 8-1 and 8-3). As another example, Berg and Purcell (1977) concluded that the relative error in determining concentration, C_0, of a substance with diffusion constant D using a receptor that effectively sampled a volume of radius r in a time t is $\Delta C_{rms}/C_0 = 0.437\ (\mathrm{D}\,t\,C_0\,r)^{-1/2}$, which also involves the reciprocal square root of time and intensity (C_0).

In Section 3-2, it was pointed out that the number of distinguishable intensities, N_i, in a receptor cell is an important determinant of its information capacity. This number can be related to intensity and noise in the following way (Snyder et al. 1977).

The stimulus that maximizes information transmission is random in nature. Assume it has a standard deviation of S_{SD}. Then if a receptor cell

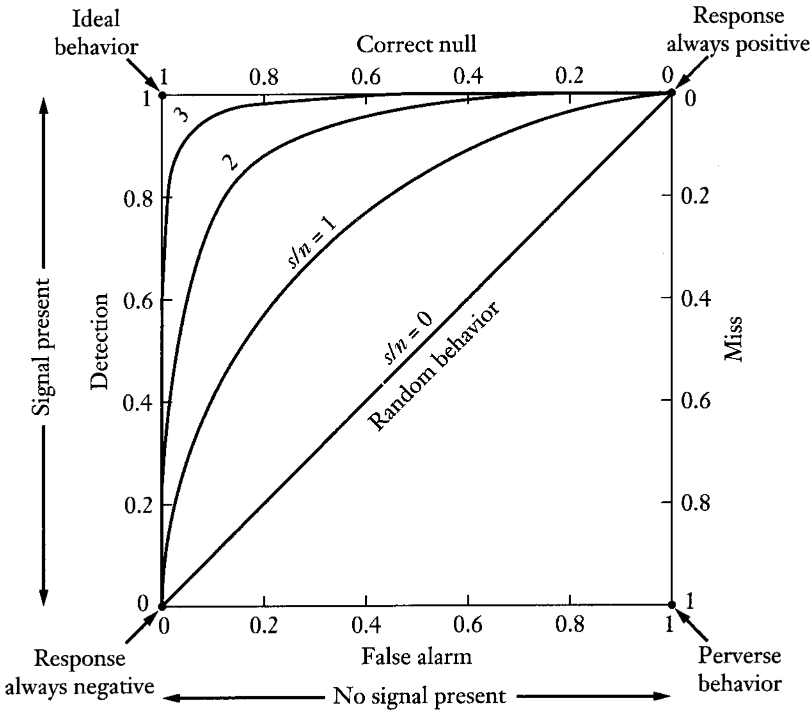

Figure 5-2 Trade-offs in different types of errors during signal detection. The axes are conditional probabilities. For example the bottom axis is the probability of a false alarm occurring, given the condition that no signal is present. For any particular signal-to-noise ratio (S/N) the probability of error depends on the threshold employed, which determines the bias of the detector for reporting either "signal" or "no signal." A detector that is given a particular S/N, which is often determined externally, can select a threshold that would produce errors corresponding to any point on the curve appropriate to that signal-to-noise ratio (0, 1, 2, and 3 are shown). The most appropriate point to select on the curve depends on the relative costs of the different types of errors. If the costs are equal, the best performance will be on the part of the curve closest to the upper-left-hand corner, which gives equal error probabilities whether or not a signal is actually present. A reduction in errors for both conditions of signal and no signal can be achieved only with an increase in the S/N ratio. In this model a normal distribution of noise was assumed, with the same width whether or not a signal was present, as in the previous figure; other distributions would change the shapes of the curves. The signal-to-noise ratio used here was determined by using the difference in mean intensity with and without the signal as the signal and the standard deviation of the intensity as the noise. With $S/N = 0$ the performance is 50 percent correct; with $S/N = 1$ the best possible performance is about 70 percent correct, and with $S/N = 3$ accuracy of 90 percent is possible.

has a noise level of N_{SD}, one can estimate the number of distinguishable intensity levels there will be. Assume that for them to be distinguished two intensity levels must differ by $2N_{SD}$, corresponding to approximately 95 percent reliability. The range of intensities is approximately twice the standard deviation of signal plus noise. Thus, dividing this by the required difference gives the number of distinguishable intensity levels:

$$N_i \cong \frac{2(S_{SD}^2 + N_{SD}^2)^{1/2}}{2N_{SD}} = \left(1 + \frac{S_{SD}^2}{N_{SD}^2}\right)^{1/2} \qquad (5\text{-}3)$$

Thus, it can be seen that the number of distinguishable levels is closely related to the ratio of the variances of the signal and the noise.

5-3 SIGNAL PROCESSING

The previous analysis assumed that noise and signal intensities and distributions were fixed. In real life, noise intensity is likely to vary, and signal intensity almost always varies with its distance from the source. Another consideration is that information concerning changes in the presence or absence of a signal is usually more important than information concerning its continued presence at a constant level. Organisms almost universally adjust to these conditions, by using a process of sensory adaptation in which changes in intensity are enhanced and steady intensities are suppressed. An example is shown in Figure 5-3.

Similar considerations apply to variations of intensity with position as well as in relation to time. One of the first steps in visual-signal processing by the retina is to enhance differences in intensity with respect to position. This general process might be called **contrast enhancement**. The following discussion of signal processing usually refers to temporal changes, but the same considerations could just as well apply to spatial patterns.

A simple model of adaptation is the process of mathematical differentiation, in which the rate of change in intensity becomes the transformed stimulus (Foster and Smyth 1980). In practice, differentiation can be accomplished simply by taking the difference between intensities at points differing in time or space by an appropriate amount. The basic requirement is for some simple form of memory and a mechanism of comparison. The differentiation, which can be performed by a simple electrical circuit of resistors and capacitors, can certainly also be done by biological cells.

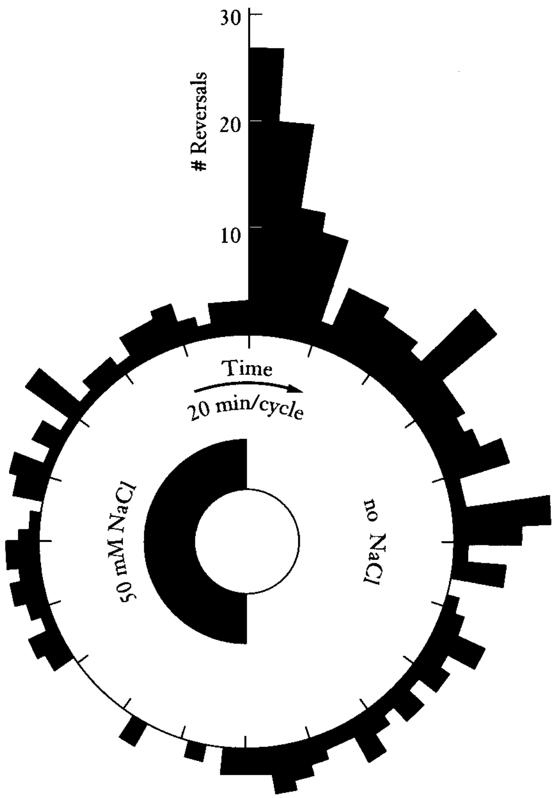

Figure 5-3 An example of sensory adaptation. Individual nematodes were exposed to two concentrations of an attractive chemical, NaCl, in an alternating cycle of 20 min, and the number of "reversal bouts," which cause a nematode to change its direction of locomotion, was recorded for each fifteen-second interval. It can be seen that immediately after the decrease in concentration there is a large increase in the probability of a reversal bout, a probability that then declines rapidly in the first minute, reaching the basal level in a few minutes as the animals adapt to the new concentration. After an increase in concentration there is a reduction in probability that lasts several minutes, until the animals adapt to this concentration. Adaptation like this is important to efficient movement along a stimulus gradient (see Section 17-2). (From Dusenbery 1980b, p. 330.)

The major question here is to identify the optimal parameters. For instance, how large a time difference should be used? If it is too large, the response will be slow; if too small, the time over which the signal can be averaged will be small and thus the two samples will be noisy. In addition, using a small time interval will mean that the difference is small. The small difference, combined with the large noise in the samples, will mean that the output will be noisy. Thus, a signal-processing system must make compromises between conflicting requirements. More-sophisticated signal-processing strategies are possible. This section introduces some of the more important concepts, with a few examples.

TIME AND FREQUENCY

In order to discuss signal processing it is convenient to introduce the concept of describing a stimulus pattern in terms of the frequencies present in it. A general mathematical result is that any function that is of practical interest, such as a pattern of intensity $I(t)$ with $0 < t < t_0$, can be described by a **Fourier series** (see, for example, Ackerman 1979, 595–99), as follows:

$$I(t) = (a_0 / 2) + a_1 \cos(\omega t) + b_1 \sin(\omega t) + a_2 \cos(2\omega t) + b_2 \sin(2\omega t) + \ldots$$

or

$$I(t) = (a_0 / 2) + \sum_{n=1}^{\infty} [a_n \cos(n\omega\, t) + b_n \sin(n\omega\, t)] \tag{5-4}$$

where

$$a_n = (2/t_0) \int_0^{t_0} I(t) \cos(n\omega\, t)dt$$

$$b_n = (2/t_0) \int_0^{t_0} I(t) \sin(n\omega\, t)dt$$

$$\omega = 2\pi/t_0$$

This relation* can be understood (Jeffrey 1979, 600–603) from the fact that the cos and sin terms represent orthogonal† functions (analogous to orthogonal axes) and the coefficients a_n and b_n are projections of the

* For readers not familiar with calculus, integrals of the form $\int_a^b F(x)\, dx$ can be interpreted as the area under a plot of the function $F(x)dx$ between the values $x=a$ and $x=b$.

† $\int \sin n\, t \sin m\, t\, dt = \int \sin n\, t \cos m\, t\, dt = \int \cos n\, t \cos m\, t\, dt = 0$ for $n \neq m$.

complete function $I(t)$ on each of the orthogonal functions, which are also described as the cross-correlations between the functions.

The sum of sin and cos terms of the same frequency can be replaced with a single sin term including a **phase** parameter, φ_n:

$$I(t) = (a_0 / 2) + \sum_{n=1}^{\infty} [c_n \sin (n\omega\, t + \varphi_n)] \tag{5-5}$$

where

$$c_n = \sqrt{a_n^2 + b_n^2}$$

and

$$\varphi_n = \arctan (a_n / b_n)$$

It is frequently useful to convert stimulus patterns back and forth between the time domain, $I(t)$, and the frequency domain, c_n's and φ_n's. The same information can be contained in either description. However, some transformations, such as that made by our sense of hearing, convert $I(t)$ to a frequency spectrum consisting of c_n's, but the phase information, φ_n's, is lost (see Section 9-2). The human sense of vision has also been described in terms of Fourier series using spatial coordinates in place of time. Thus the spatial resolution of the visual system is described in terms of its sensitivity to different spatial frequencies.

The terms in the infinite series oscillate at increasingly higher frequencies, and the more terms included in the series the better the match to the original function (see Figure 5-4). This result demonstrates that an arbitrary pattern contains components of many different frequencies; in other words, it has a wide bandwidth.

FILTERING

A basic form of signal processing is a **filter** that continuously transforms a signal from one form to another. The simplest filters, which remove high or low frequencies from the signal, are called **low-pass** or **high-pass filters**

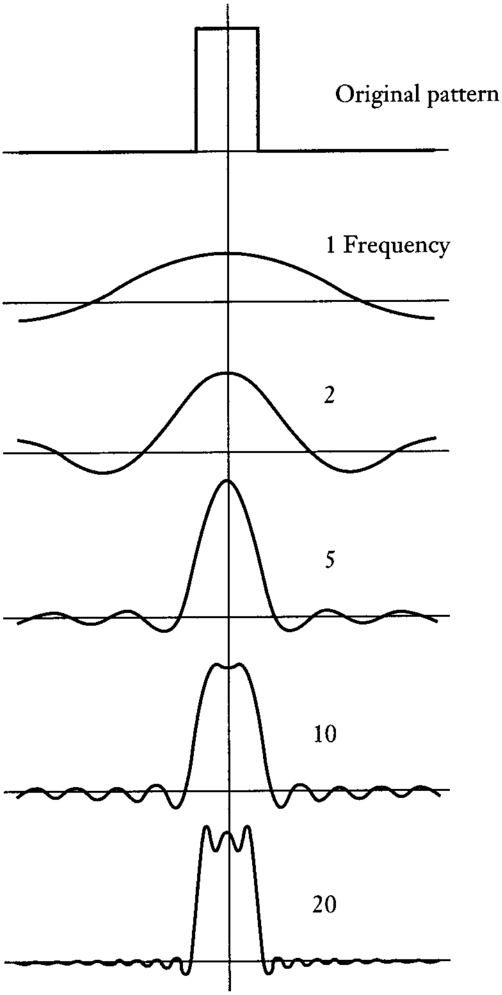

Figure 5-4 An example of Fourier frequencies summing to match a pattern. The terms in the Fourier series were weighted to sum up to the rectangular pulse shown at the top. The lower curves show the shapes produced when increasing numbers of terms are included in the summation; the first curve includes only the first term, the bottom curve the first twenty terms. The curves approximate the rectangular pulse more and more accurately as more terms are employed. This comparison indicates how infinite series can lead to an accurate representation of the original rectangular pulse and shows how the first few terms can be used as an approximation for some purposes.

respectively. All real detectors have an upper frequency limit that is determined by how fast they can respond to a change in input. Consequently, they act as low-pass filters. Low-pass filtering is desirable to the extent that it effectively increases the time interval over which signals are averaged, which can be beneficial in increasing the signal-to-noise ratio, as discussed in the previous section, but it has the disadvantage that response is slowed.

If the signal of interest is known to have components that exist in a narrow band of frequencies, a **band-pass filter** that reduces the frequency components outside the band can improve the signal-to-noise ratio. For instance, crickets have receptors that are tuned to the frequency of the song of their species (Hutchings and Lewis 1983). Specifically, the calling song of *Gryllus campestris* has a carrier frequency of 4 kHz, and this species is most sensitive to sounds at this frequency. At other frequencies, sounds must be more intense to be detected. Consequently, noise at other frequencies has less impact on detection of the calling song. In the simplest case, when noise consists of completely random changes in intensity all frequencies will be present equally. (This situation is called white noise, in analogy with the fact that the color white contains all colors represented equally.) A cricket exposed to white noise filters out all frequencies except those in a narrow band. Doing so reduces the noise intensity, improving the signal-to-noise ratio.

A simple example of a low-pass filter is a **running average**, in which intensities from a local interval are averaged together to produce the output. This interval then slides along the intensity pattern to produce the output pattern. All real detectors act this way to some extent, because of their limited frequency response. A running average can be elaborated into a band-pass filter by making the output the difference between the averages of two adjacent intervals. In other words, in this case the filter has both positive and negative lobes. (These effects can be seen below in Figures 5-6 and 5-8.)

The behavior of a signal-processing mechanism that is linear* can be completely predicted by knowing its response, $R_i(t)$, to an impulse signal occurring at $t = 0$, which consists of a pulse of sufficiently short duration that the response depends on only the product of the amplitude and duration of the pulse. The response, $R(t)$, to any stimulus pattern, $I(t)$, can be represented mathematically by the convolution integral of the stimulus and the impulse response (Hancock 1961, 126):

* Linear means that filtering the sum of two different inputs produces an output that is the same as the sum of the outputs of the two individually filtered inputs. Real filters are never completely linear, because they have limited response ranges. However, many real filters approximate linearity for small signals, which are usually the most important ones.

$$R(t) = \int_{-\infty}^{t} I(\tau)\, R_i(t-\tau)\, d\tau \tag{5-6}$$

This integral simply represents a running average with the possibility that it can be formed by weighting various parts of the interval differently, according to the shape of $R_i(t)$. For a normal running average $R_i(t)$ would have a constant value within a certain time interval and be zero outside that interval.

Another way of understanding this integral is to consider that the stimulus pattern $I(t)$ can be represented by a series of impulses, as shown in Figure 5-5. Since the filter is assumed to be linear, its output at any given time is simply the sum of the remaining responses to each of the previous impulses. Practical filters have impulse responses that decay to zero over some time interval t_i, and only a stimulus intensity this far into the past will influence the output at time t. Thus, the limits of integration could be taken as $t - t_i$ to t, rather than as the infinite limits given in equation 5-6.

If the stimulus is composed of the sum of a signal $S(t)$ to be detected and noise $N(t)$, the output can be represented as the sum of two integrals involving the signal and noise independently:

$$R(t) = \int_{-\infty}^{t} S(\tau)\, R_i(t-\tau)\, d\tau + \int_{-\infty}^{t} N(\tau)\, R_i(t-\tau)\, d\tau \tag{5-7}$$

The optimum technique for detection, a **matched filter**, is obtained when the impulse response matches the signal to be detected:

$$R_i(t) = S(t_0 - t) \tag{5-8}$$

where the signal occurs at time t_0. Thus, (Schwartz 1963, 149)

$$R(t) = \int_{-\infty}^{t} S(\tau)\, S(t_0 - t + \tau)\, d\tau + \int_{-\infty}^{t} N(\tau)\, S(t_0 - t + \tau) d\tau \tag{5-9}$$

The first integral of the pair can be interpreted as comparing a copy of the expected signal shifted in time to the actual signal. At time $t = t_0$ there will be a match, at which point both terms within the integral will have the same sign for all values of τ and will be large for the same values of τ. Consequently, their product will always be positive and the integral will sum to a large value. In contrast, the second integral involves the noise, which would generally not

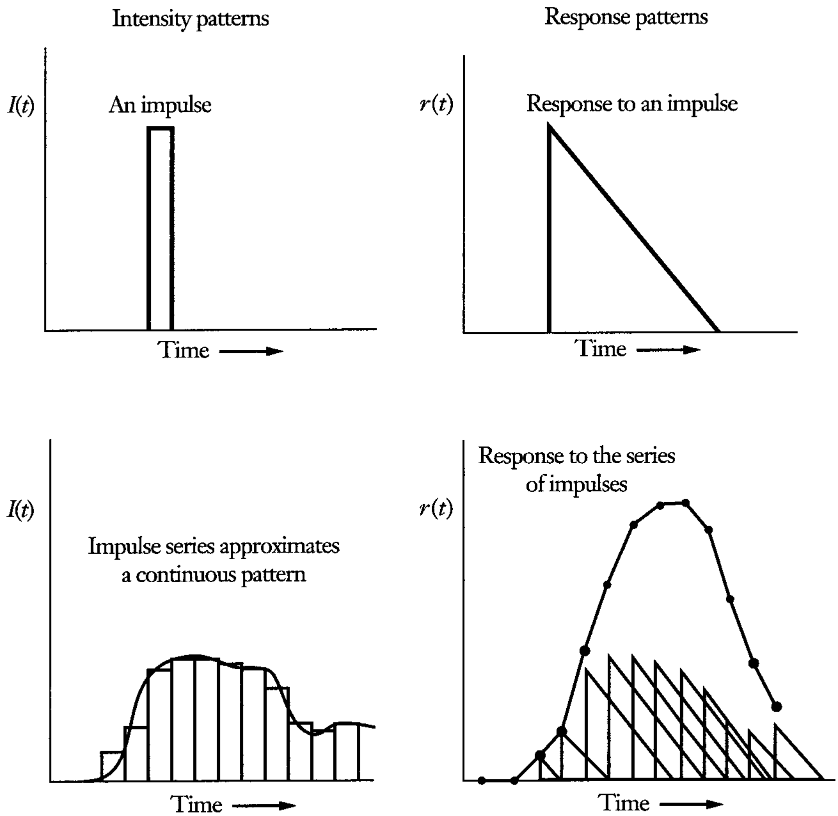

Figure 5-5 A graphical representation of an impulse response and its convolution. *Left*: Two intensity patterns. *Right*: Responses to these patterns. The top pair of graphs shows the intensity pattern of a single impulse represented by the rectangle and the response to it, in this case a triangle with instantaneous rise and decay at a fixed rate assumed. The bottom graphs show that a continuous intensity pattern can be approximated by a series of impulses of the appropriate size, and then (assuming that the receiver is linear) the response to the pattern can be accurately determined at any time simply by summing the remaining effect of each previous impulse. This is the procedure carried out by the convolution integral of the stimulus and impulse response in equation 5-6.

be correlated with the expected signal. The two terms in the product will often have opposite signs, and one will be small when the other is large. Thus, the second integral will generally sum to a small value. In this way the processed signal will be enhanced when the input stimulus contains a pattern resembling the signal; for example, see Figure 5-6.

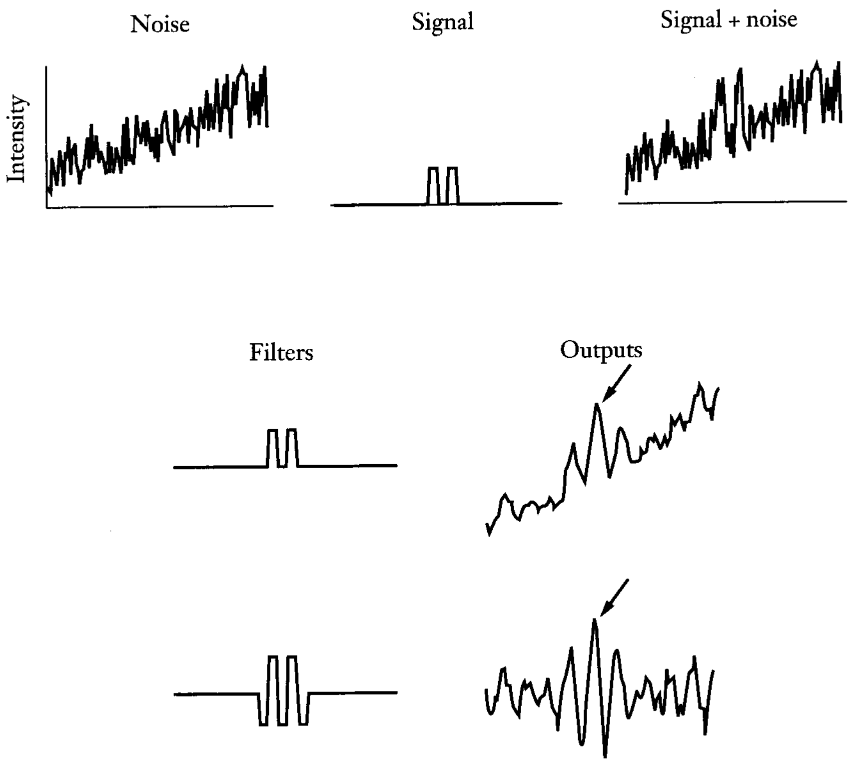

Figure 5-6 Examples of filters matched to the signal. *Top*: The noise consists of 100 random numbers ($0 \leftrightarrow 2$), representing high-frequency noise, superimposed on a drift ($0 \rightarrow 2$), representing low-frequency noise. The signal to be detected consists of two pulses ($0 \leftrightarrow 1$). The intensity received, that is, the objective intensity, is the sum of the signal and the noise. In this example the signal is marginally discernible to the eye. A filter with two positive lobes matching the signal (*middle row*) produces an output that enhances the signal but increases with the drift in intensity. A similar filter, with equal positive and negative lobes (*bottom row*) enhances the signal and eliminates the drift in intensity. The averaging within each lobe reduces the high-frequency noise, while taking the difference between adjacent intervals with equal weight removes the drift from low-frequency noise.

This strategy can be illustrated with bacterial chemotaxis (see Section 17-2), where the impulse response has been measured (Block et al. 1982). It has a positive peak lasting about 1 s, followed by a negative peak of about 3 s duration (Figure 5-7, upper right). The areas of both peaks are ap-

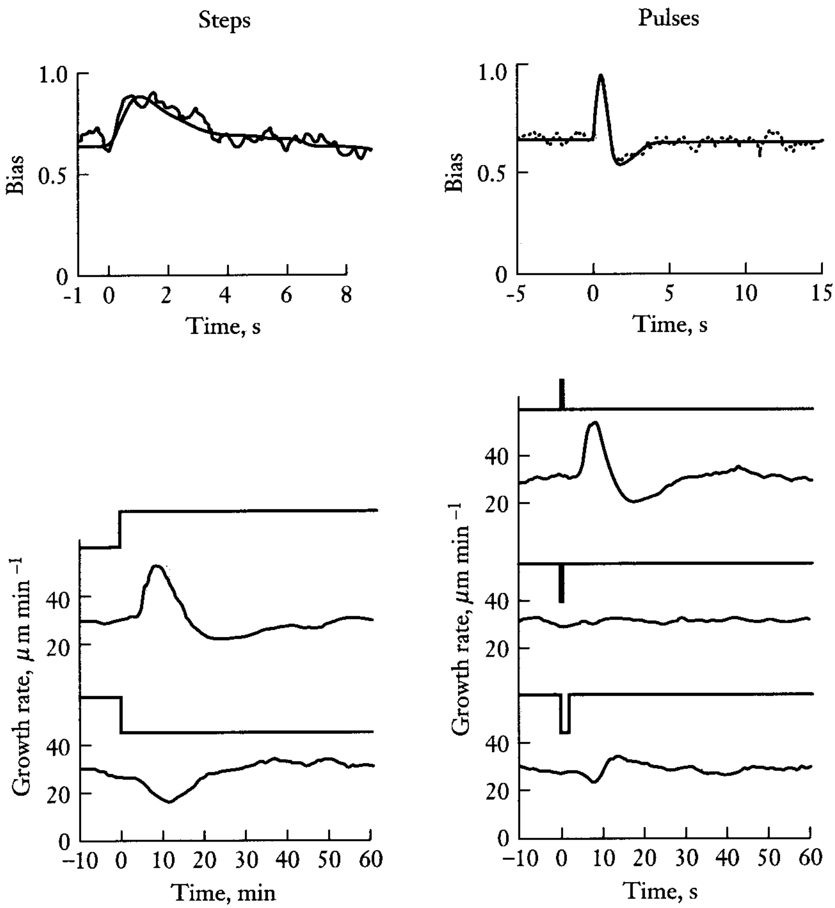

Figure 5-7 Stimulus response kinetics of bacterial chemotaxis and mold phototropism. Behavioral responses of two very different microorganisms to step increases in intensity (*left*) and to brief pulses (*right*). *Top*: The response is the bias in the direction of turning of the flagella of cells of the bacterium *Escherichia coli*, and the stimulus is the concentration of an amino acid (Segall 1986). *Bottom*: The response is the rate of growth of sporangiophores of the mold *Phycomyces blakesleeanus*, and the stimulus is light intensity (Galland 1987). Both organisms respond in a manner suggestive of the filter in Figure 5-8.

proximately equal. This response pattern indicates that the signal-processing mechanism sums up the stimulation over the past 1 s which reduces noise, and compares it by subtraction to the average stimulation over the previous 3 s, which provides sensitivity to change. The equality of the positive and negative peaks means there is no response to a constant stimulus. The Fourier transform of the impulse response is a band-pass peaking at 0.25 Hz.

A similar pattern of signal processing is demonstrated in a mold. Sporangiophores of *Phycomycetes blakesleeanus* grow toward light, a phototropism that has been studied in great detail (Galland and Lipson 1987; Presti and Galland 1987). Of interest here is the fact that the kinetic pattern of the response to a change in intensity is similar to that of bacteria, but with a slower time scale (see Figure 5-7, bottom). This type of filter is probably a good match for many stimulus situations, because it provides band-pass characteristics to reduce noise and emphasizes changes in intensity that usually carry more-important information than do steady levels of intensity (Foster and Smyth 1980). Its effects are illustrated in Figure 5-8.

For unknown signal patterns or general signal processing the optimum filter design is not clear. A general property of Fourier transforms is that there is an inverse relationship between the spread of a function in time or space and the width of its Fourier transform in the frequency domain (bandwidth). For some examples see Figure 5-9. Furthermore, the bandwidth of a pulse of constant frequency is approximately the reciprocal of its duration (Pye 1983).

Marr and Hildreth (1980) have argued that the optimal spatial filter for vision should be one having minimum width in both the spatial and frequency domains. The function that uniquely satisfies these requirements is the Gaussian function. To use this function as a detection device, take the second derivative and define detection as occurring when the output crosses zero. This process leads to a function that has a "Mexican hat" shape (Figure 5-10) that is similar to a difference in two Gaussian distributions (Ratliff 1978) with widths in the ratio of 1.6 : 1. The effect of convoluting this pattern with rectangular pulse signals is shown in Figure 5-10.

Another basic question concerning detection is to determine at what distance from the target it occurs—in other words, what is the range of detection? As discussed, this depends on the output of the source, the attenuation in the environment, the noise present, and the characteristics of the receiver. The general aspects of attenuation were discussed in Chapter 4, on signal transmission. More specific aspects of all three considerations are discussed in the chapters on specific stimuli in Part 2.

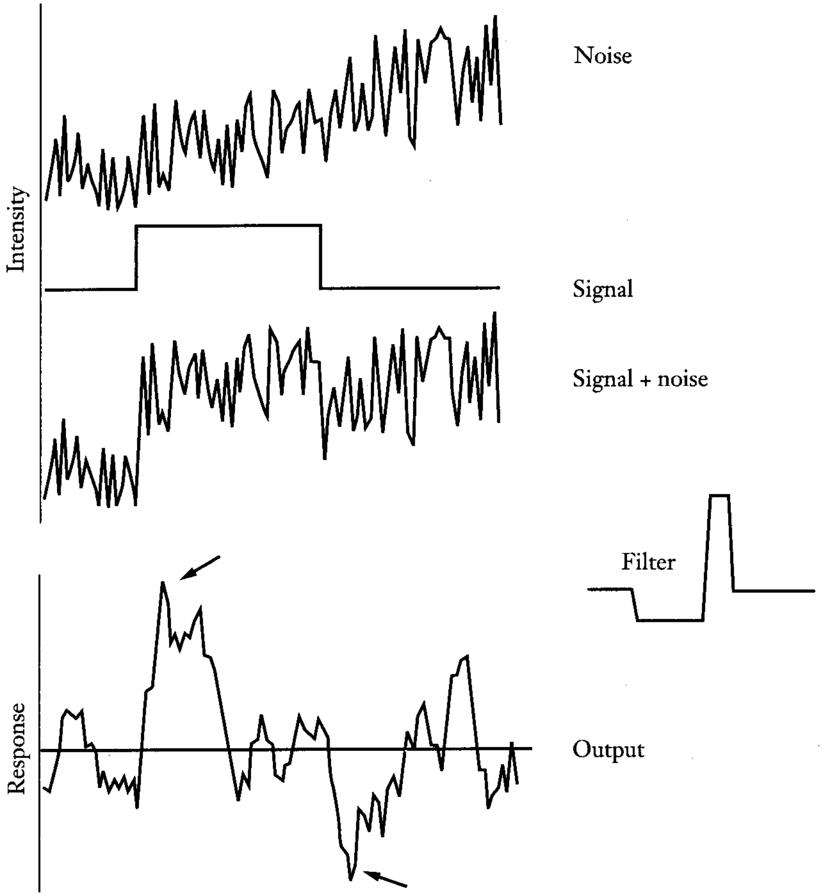

Figure 5-8 The effect of an asymmetrical filter. The noise (*top*) is the same as in Figure 5-6. The signal is a single pulse (0 ↔ 1). The intensity received (the objective intensity) is the sum of the signal and the noise. In this example the signal is marginally discernible to the eye. The input is filtered by taking three times the average of five adjacent points (positive lobe of filter) and subtracting the average of fifteen points from an adjacent interval (negative lobe of filter). This filtering pattern, which is appropriate for many organisms (see Figure 5-7), yields the output response (or subjective intensity) (*bottom*). Averaging within each lobe reduces high-frequency fluctuations and taking the difference between the two lobes of equal weight removes the drift of low-frequency noise. More importantly, the changes in the signal are represented by a positive peak (*top arrow*) for the increase in signal intensity and a negative peak (*bottom arrow*) for the decrease in signal intensity. Thus, contrast information is emphasized and changes in intensity can be easily detected as levels of the output response.

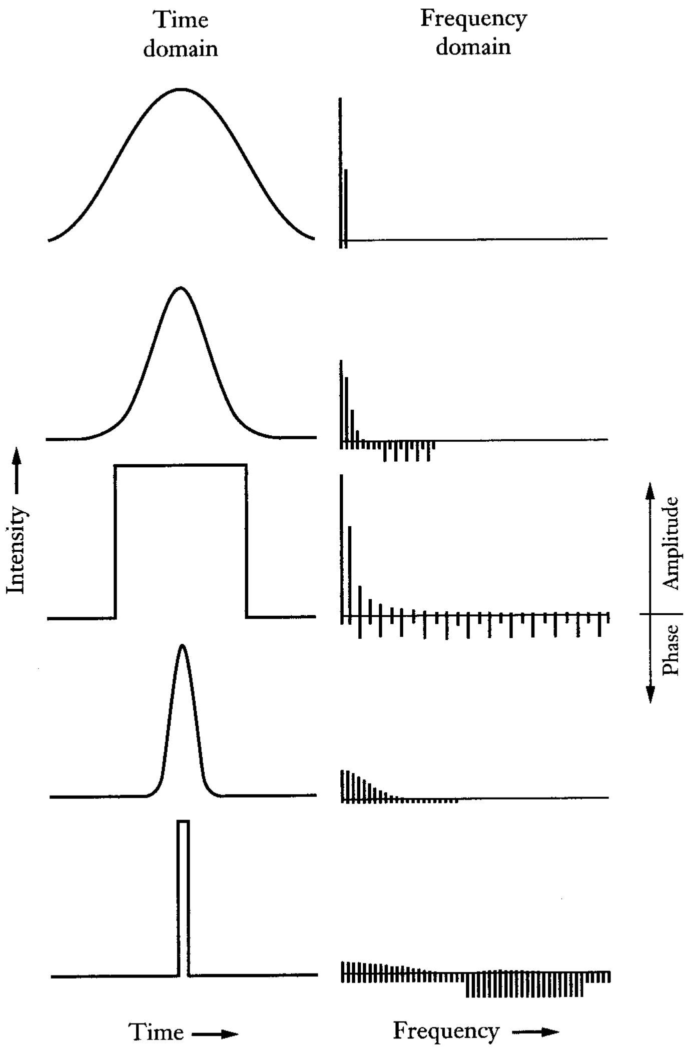

Figure 5-9 Relationships between time and frequency representations. Each row contains a graphical representation of the same pattern, presented in the time domain on the left and the frequency domain on the right. In the frequency domain the intensity of each frequency component is indicated by the height of the line above the axis, with the phase indicated below the axis. The frequency plot starts with 0 frequency (*left*) and extends to fifty times the reciprocal of the period in the time domain. A wide pulse in the time domain has a narrow frequency distribution (*top*), while a pulse of short duration has a wide range of frequency components (*below*). Rapid changes in intensity, as in rectangular pulses, are associated with high-frequency components.

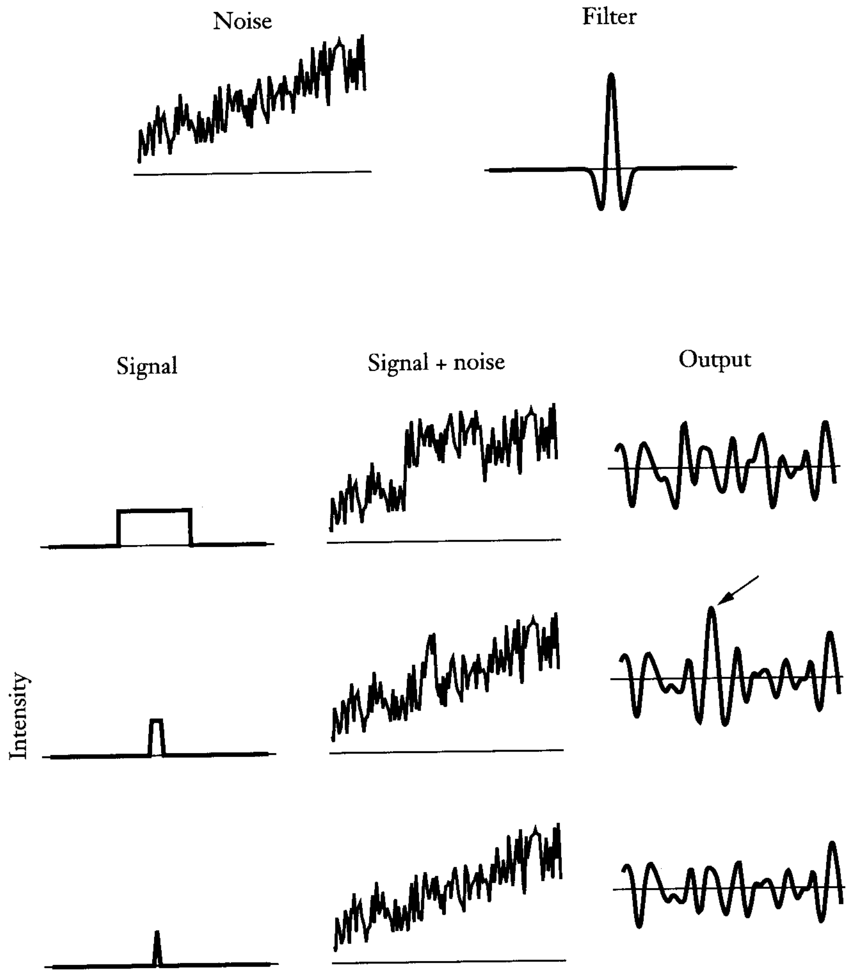

Figure 5-10 Difference of Gaussians filter. The noise pattern is the same as in previous figures. The filter is the difference between two Gaussian distributions, which produces the "Mexican hat" shape. The signals are single pulses of the same height but of three different durations. This filter yields a strong response only to a signal of the appropriate width (*arrow*).