

Figure 5–1 Hippocampal place cells and place cell maps.

A. Input—output transformations occur in the trisynaptic circuitry of the mammalian hippocampus, proceeding from the dentate gyrus input region, to the CA3 area, and to the CA1 output region, with principal excitatory neurons (red) in each region as primary processing units. Activity of principal cells is modulated by local circuit GABAergic interneurons (gray).

B. Place cell firing in the hippocampus. The path taken by a rat is shown in **black** as it traverses a square arena. Electrodes

the hippocampus of several other mammalian species, including bats, monkeys, and humans. Distinct sets of place cells are activated by distinct locations in a given environment. Consequently, although individual place cells represent relatively small spatial areas, the full diverse population of place cells in the hippocampus tiles the entire environment, and any given location is encoded by a unique ensemble of cells. The hippocampal place coding network provides an example of a cognitive map, initially postulated by the psychologist Edward Tolman, that enables an animal to successfully remember and then navigate its environment. The role

were implanted within the hippocampus to record from individual cells. Above: A single place cell increases firing (each action potential represented by a red dot) at discrete locations in the environment. Below: A color-coded heat map of firing frequency of the schematic place cell. Lower wavelength colors (yellow and red) represent higher firing rates on a background of no activity (dark blue).

C. Color-coded heat maps showing the firing of 25 different place cells recorded simultaneously in the hippocampal CA1 region as the rat explores a square box.

of the hippocampus in memory formation and the mechanisms by which the hippocampal spatial map is encoded are explored in detail in Chapters 52 and 54.

The electrophysiological methods available to O'Keefe in 1971 were limited to recording one place cell at a time, but subsequent advances allowed investigators to record dozens, and more recently hundreds, of place cells simultaneously. Critically, while single place cells encode only specific parts of the environment and are prone to occasional noisy firing outside of their place fields, entire populations of place cells provide more complete spatial coverage and the

reliability of redundant place coding. These features of population coding have paved the way for new and powerful computational analyses. In particular, it is possible to decode the activity of populations of place cells and estimate an animal's location within an environment. This is accomplished by determining each cell's spatial selectivity and using this selectivity as a template to decode ongoing activity. In practice, this decoding is often performed by weighting each cell's contribution to the final estimate of the animal's position by a factor proportional to that cell's spatial coding reliability. Using this and similar techniques, one can reconstruct an animal's location from second to

second within room-sized environments with a precision of a few centimeters (Figure 5–1C).

Hippocampal function has been strongly implicated in spatial and declarative memory based on studies using spatial decoding techniques. During active exploration of an environment, hippocampal activity reflects place coding, but during immobile or resting behavior, the hippocampus enters a different regime in which neural activity is instead dominated by discrete semi-synchronous population bursts termed sharp-wave ripples (Figure 5–2A). These events are thought to be internally generated by circuitry within the hippocampus.

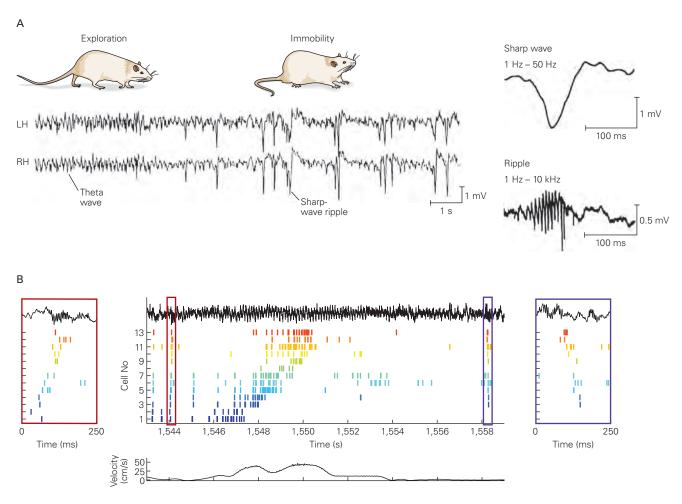


Figure 5–2 Hippocampal sharp-wave ripples and sequence replay.

A. Left: Behavior dependence of hippocampal local field potential activity (LH and RH, left and right hippocampus). Theta waves are present during exploration, and large negative sharp waves during immobility. Right: Sharp waves and ripples recorded from the hippocampal CA1 region. (Adapted, with permission, from Buzsaki 2015; and reproduced, with permission, from Buzsaki et al. 1992. Copyright © 1992 AAAS.)

B. Place cell sequences experienced during behavior (*middle*) are replayed in both forward (*left*) and reverse (*right*) direction during sharp-wave ripples. The rat moved from left to right on a familiar track. Spike trains for place fields of 13 CA3 pyramidal cells while the rat is on the track are shown before (forward replay; red box), during (*middle*), and after (reverse replay; blue box) a single traversal. The CA1 local field potential is shown on top (black traces), and the animal's velocity is shown below. (Adapted, with permission, from Diba and Buzsaki 2007. Copyright © 2007 Springer Nature.)

Notably, sharp-wave ripples are prominent during resting periods after recent learning, for example after exploration of an environment. Spatial decoding of the activity of place cells active within these short (50 to 500 ms) sharp-wave ripples reveals that hippocampal neurons recapitulate or replay discrete trajectories through the recently explored environment. Although these trajectories replicate paths taken through space, the replayed activity sequences differ from those observed during active exploration in several ways.

First, replayed sequences within sharp-wave ripples are time compressed, occurring about 10 to 20 times faster than during exploration (Figure 5–2B). Second, they can occur either in the same direction as behavioral spatial trajectories (forward replay) or in the opposite direction (reverse replay). Thus, decoding a single postexploration, 200-ms, sharp-wave ripplereplay event may reveal a virtual mental trajectory spanning 2 to 4 seconds of behavioral time replayed backward from how it was experienced. Replay is thought to represent a form of mental rehearsal by which certain memories are gradually consolidated and thus may be a crucial aspect of the role of the hippocampus in memory.

Neural Circuit Motifs Provide a Basic Logic for Information Processing

Neurons tend to be highly interconnected, both with nearby neurons and with neurons in distal brain areas. Knowledge of neuronal connections, called connectomics, is expanding rapidly due to a number of new methods for uncovering fine-scale anatomical structure. Patterns of neuronal interconnection come in several varieties.

Connections from one area to another, for example from the thalamus to primary visual cortex, are termed feedforward (Figure 5-3A). The forward direction is defined as extending from a more peripheral or primary area, such as the retina, thalamus, or primary visual cortex, to a higher area with more complex response properties, such as the visual areas that respond selectively to particular objects. In most cases, two areas that have feedforward connections also have feedback connections; for example, there are numerous connections from primary visual cortex back to the thalamus. Local connections often extend from one neuron to another, ultimately looping back onto the original neuron. Such looping connectivity is called recurrent. Many neurons are involved in all of these types of connectivityfeedforward, feedback, and recurrent—but it is useful to consider the functional implications of these different connectivity motifs separately.

Connections between neurons can be either excitatory or inhibitory. Normally, excitatory connections lead to increased neural firing and inhibitory connections lead to decreased neural firing. Many neural circuits receive strong excitatory drive from hundreds or thousands of synapses. If not checked by inhibition, this synaptic excitation would lead to unstable neural activity. A near balance of excitation and inhibition is a common feature of neural circuits that may enhance their computational capacity. However, this fine tuning may make the circuits prone to generating seizure activity if the balance between excitation and inhibition is not properly maintained, as occurs during epilepsy.

In mammals, visual information is processed in a series of brain areas that are often approximated as having feedforward circuitry. Feedforward circuits can process information in sophisticated ways, for example

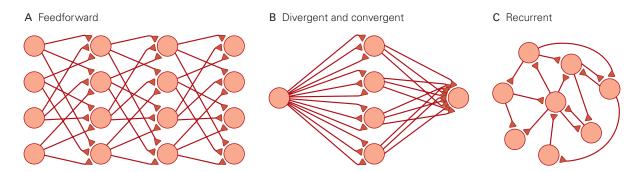


Figure 5-3 Four basic neural circuit motifs.

A. A feedforward circuit in which synaptic connections extend in a single direction from one processing level of neurons to another.

B. Divergent feedforward connections describe a small number of presynaptic neurons connecting to a larger number.

Convergent connections describe a large number of presynaptic neurons connecting to a smaller number.

C. In a recurrent network, synaptic connections occur in multiple directions between neurons, forming looping pathways through the circuit.

extracting and identifying objects from a complex visual scene, but they cannot produce ongoing, dynamic patterns of activity. For this purpose, recurrent circuitry is needed (Figure 5–3C).

Within feedforward circuitry, two submotifs can be identified: divergent and convergent connections (Figure 5–3B). In divergent connections, the number of neurons that receive a given type of input exceeds the number of neurons providing that input, so the information encoded in the presynaptic input neurons is expanded in the postsynaptic output neurons. In convergent connections, many presynaptic neurons send input to a smaller number of postsynaptic neurons.

The most prominent example of both divergent and convergent connectivity is provided by the cerebellum, as discussed later.

Visual Processing and Object Recognition Depend on a Hierarchy of Feed-Forward Representations

Visual information is processed within a large number of brain regions arranged hierarchically (Figure 5–4). Moving up the hierarchy from the primary sensory input generated by the retina, neurons respond to increasingly complex combinations of visual features, culminating in selectivity for complex objects, such as

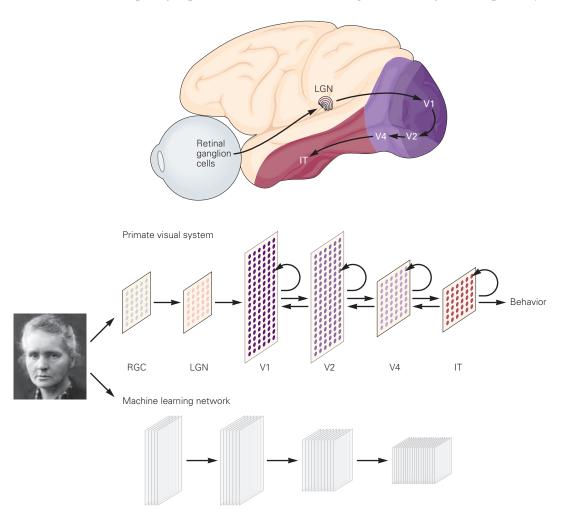


Figure 5-4 Comparison of biological and machine learning networks. In the visual system, multiple brain regions form a hierarchy in which neurons in series become progressively selective to more complex objects. The regions in the primate visual system pathway represent retinal ganglion cells (RGC), the lateral geniculate nucleus (LGN) of the thalamus, ventral stream visual areas (V1, V2, and V4), and the inferotemporal cortex (IT). The number of neurons per region varies (represented by the colored dots), but their selectivity steadily

increases. The machine learning network pathway represents layers of a feedforward network trained to identify objects in images. Increased selectivity in the different regions of the machine learning network is indicated by the growing numbers of stacked sublayers, reflecting selectivity to a richer array of visual features. The hierarchy of response selectivities recorded in different visual areas resembles the activities seen in corresponding layers of the machine learning network. (Adapted, with permission, from Schrimpf et al. 2018.)

faces. Considerable research is devoted to identifying principles upon which the structure of the visual hierarchy is based. The development of artificial neural network models in machine vision has proven to be an instructive analogy for addressing this issue.

From the retina, to the thalamus, to the primary visual cortex, onto the highest visual areas associated with cognition in inferotemporal cortex, visual neurons respond selectively to particular patterns of light, dark, and color in regions of the visual field called their receptive fields. From the lowest to highest stages of visual processing, neurons have increasingly larger receptive fields and higher degrees of selectivity. At each stage, neurons with a particular type of selectivity tend to have receptive fields that tile the visual scene, providing full coverage for the selected feature. Moreover, the arrangement of the receptive fields in each visual brain area is topographically matched to the layout of the image of the external world on the retina, that is, the cortex forms a map of the visual field.

As receptive fields enlarge and selectivity increases, neural responses depend less on the precise location of the selected object or pattern and more on its overall features. In general, neurons in higher stages of visual processing respond more selectively to a larger portion of the visual field and depend less on features such as location, size, and orientation. This correlates with our ability to recognize objects independent of their location, size, and orientation in a scene. At the highest stages of the hierarchy, neurons can, for example, respond selectively to particular faces located across

the visual field, independent of the size of the face or its angular pose (ie, head direction).

The ideas of tiling, increased receptive field size, increased selectivity, and decreased dependence on view-dependent factors are central to the construction of artificial networks for machine vision. Such networks can reach human-level performance on some object recognition tasks. Furthermore, the pattern of errors that the machines make on difficult images matches, to some degree, the errors made by human subjects. Nonhuman primates can also perform these tasks at levels comparable to humans, and interestingly, recordings from different visual areas along the object recognition pathway correspond to activity seen in the artificial networks at similar stages in visual processing (Figure 5–4).

Diverse Neuronal Representations in the Cerebellum Provide a Basis for Learning

The most abundant class of neurons in our brains are the roughly 50 billion granule cells at the input stage of the cerebellum, composing more than half of all the neurons in the brain. The cerebellum is a hindbrain structure vital for motor coordination but also implicated in the adaptive regulation of autonomic, sensory, and cognitive functions (Figure 5–5). Malfunction of cerebellar circuits may contribute to various neurological disorders, including autism. In contrast to the thousands of inputs that most brain neurons receive, each granule cell receives just a handful of inputs (four on average).

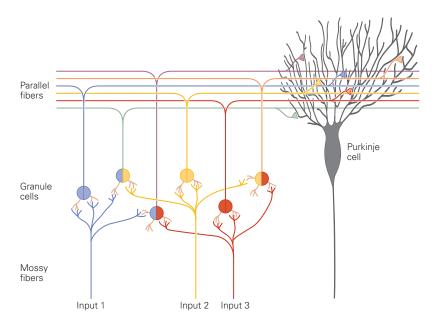


Figure 5–5 The cerebellum receives input from many regions of the brain and spinal cord. These inputs, known collectively as mossy fibers, are recoded in a vast number of granule cells, an example of divergent connectivity, allowing for many possible mixtures of the input signals. Dendrites of Purkinje cells receive convergent input from hundreds of thousands of granule cells relayed by their axons, known as parallel fibers. Parallel fiber to Purkinje cell synapses are modifiable, which is believed to be an important mechanism underlying motor and possibly other forms of learning.

Recent experimental findings using neuroanatomical tracing and electrophysiological recording indicate that inputs converging onto a single granule cell often originate from distinct brain regions. As a result, the firing of individual granule cells may represent any one of an enormous range of combinations of stimuli or events. For example, a cell may fire only during the conjunction of a specific visual stimulus (such as a moving tennis ball) with the movement of a particular body part (such as the flexing of the wrist). Representations that combine different types of information in this way are called mixed.

Cerebellar granule cells provide an extreme example of divergent feedforward connectivity, with the information carried by approximately 200 million input fibers (called mossy fibers) mixed and expanded onto the 50 billion granule cells. Such a large representation is needed to handle the many different ways that multiple channels of information can be combined. For example, representing all possible combinations of 2 out of just 100 different input channels requires $100 \times 99/2$, or 4,950, different response types. Requiring a representation of all triplets pushes this number up over 150,000, and the number increases rapidly for four and more combinations. Because the large number of possible combinations would be difficult to specify genetically, it is generally thought that the assignment of mossy fibers to their granule cell targets is largely random.

This analysis suggests that the role of the cerebellar granule cells is to combine a large number of input channels in many possible ways. Such a representation clearly would be useful for making inferences and generating actions that depend on the co-occurrence of combinations of stimuli and actions. However, to be useful, this information must somehow be read out from the huge number of granule cells.

Read-out from the cerebellar cells is accomplished by Purkinje cells, the output neurons of the cerebellar cortex. In contrast to the highly divergent connectivity at the inputs to granule cells, connections between granule cells and Purkinje cells provide an extreme example of convergence. A single Purkinje cell receives input from over a hundred thousand granule cells. Theories of cerebellar function developed in the 1970s by David Marr and James Albus posited that this convergence allows Purkinje cells to extract useful information from the extremely rich representation provided by granule cells. By doing this, Purkinje cells may, for example, underlie the amazing human capacity to form the many complex associations required for motor skills, such as riding a bicycle or playing a musical instrument. However, to extract information that is useful for a number of purposes under a variety of conditions, the read-out provided by Purkinje cells must be adaptable. This adaptability is provided by the plasticity of the synapse between a granule cell and Purkinje cell synapse, as discussed in a later section.

Recurrent Circuitry Underlies Sustained Activity and Integration

Neurons are inherently forgetful. Transient synaptic input typically evokes a brief response that decays within a few tens of milliseconds. The time course of this decay is determined by an intrinsic property of neurons known as the membrane time constant (Chapter 9). How then do patterns of neural activity persist long enough to support cognitive operations such as memory or decision making that play out over seconds, minutes, or even longer periods of time?

Consider, for example, trying to detect whether you hear a familiar voice in a crowded room full of people talking loudly. As you listen, you may occasionally detect a bit of sound that resembles the voice you are searching for but that by itself is inconclusive. Nevertheless, over time, you may be able to accumulate enough evidence to arrive at a conclusion. This process of evidence accumulation requires integration, meaning that a running sum must be maintained and augmented as additional evidence is detected. Integration requires both a computation (addition) and memory to compute and maintain a running total (Chapter 56).

For a neural circuit to perform integration, a transient input must produce activity that is sustained at a constant level even after the input is gone. This sustained activity provides a memory of the transient input. As outlined in the previous paragraph, circuits that integrate can be useful for accumulating information, but they are also needed for noncognitive tasks such as maintaining the constant muscle tension required to hold a fixed body posture. One of the best studied neural integrators is the circuitry that allows humans and animals to maintain a constant gaze direction with their eyes, even in the dark. The fact that eye movements can be studied across a wide range of species, from fish to primates, has greatly facilitated progress. Moreover, the relative simplicity of the oculomotor system has fostered fruitful dialog between experimental and theoretical studies. (The oculomotor system is described in more detail in Chapter 35.)

The existence of integrator circuits in the oculomotor system was first suggested by a puzzling observation from neuronal recordings (Figure 5–6A). Oculomotor neurons that control the eye muscles increase action potential firing transiently to evoke

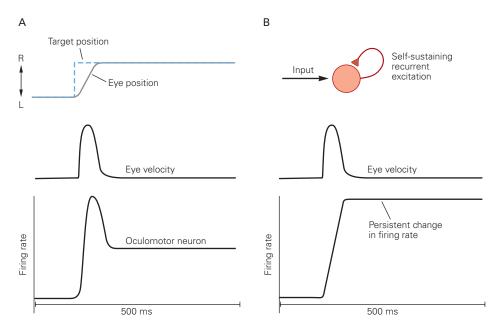


Figure 5–6 Recurrent circuitry and sustained neural activity are required for maintaining eye position.

A. Above: A saccadic eye movement consists of a rapid movement change in eye velocity to bring a target back to the center of gaze. This is followed by a sustained change in eye position to maintain the fovea on the target. The dashed blue line shows the location of the target, and the gray line shows the

eye movement and subsequent fixation on the target at its new position. *Below:* An oculomotor neuron exhibits a brief burst of activity related to eye velocity along with sustained activity related to eye position.

B. Recurrent excitation can explain how a brief pulse of input, such as an eye velocity signal, can lead to a persistent change in firing rate through a process akin to mathematical integration.

movements of the eye but also exhibit sustained action potential firing needed to hold the eye in fixed position. For example, a motor neuron that projects to an eye muscle that moves that eye to the left will fire at a high rate when gaze is maintained left of center and at a low rate when gaze is maintained right of center. The puzzle is that the premotor neurons in the superior colliculus and brain stem that project to the oculomotor neurons only fire transiently before eye movements. They do not show any sustained activity related to eye position. How then is this sustained activity generated?

An early conjecture, now strongly supported, is that steady eye position signals are computed by brain stem neurons that integrate the transient eye velocity signals. Such neurons receive velocity information and provide the steady output to the oculomotor neurons that maintain eye position. Lesions or inactivation of certain brain stem nuclei in monkeys, including the medial vestibular nucleus and the nucleus prepositus hypoglossi, result in a failure to maintain steady horizontal eye position following eye movements, suggesting that the neural integrator circuit lies within these structures. Damage to these brain stem structures in humans leads to the same problem, known clinically as gaze-evoked nystagmus (Chapter 35).

How do neural circuits perform integration? One possibility is that integration is supported by specialized intrinsic neuronal properties that effectively lengthen neuronal membrane time constants, allowing brief inputs to generate sustained output. A variety of candidate mechanisms have been described involving different voltage-activated ion channels. However, studies using intracellular recordings, which allow for direct control over the membrane voltage of the recorded neuron, have shown that sustained positionrelated signals persist even when the neuron's voltageactivated channels are blocked. A second possibility is that integration arises from interactions among a network of synaptically coupled neurons. Intracellular recordings in goldfish support this idea by showing that levels of synaptic input vary with eye position.

The question of what types of neural networks are capable of performing integration has been explored extensively in theoretical studies. One class of models that has been considered relies on recurrent connectivity, specifically a population of neurons that excite each other. A weakly coupled network of this type responds to an input pulse with activity that rapidly decays away. Increasing the strength of the recurrent excitation adds back some of the activity that would

otherwise decay, lengthening the duration of the population response. If recurrent excitation is increased to the point where the recurrent excitation set up by a transient input precisely cancels the decay, the response can last indefinitely. This requires fine-tuning of network parameters.

In a perfectly tuned network, a transient pulse of input produces a change in firing rate that lasts forever in the absence of further input. Equivalently, such a population computes a running integral of the input it receives (Figure 5–6B). If the transient excitation in the network is not perfectly tuned but instead is slightly weaker, the input produces a change in firing rate that decays slowly. Eye position, in the dark, tends to drift back to the center in about 20 seconds, suggesting that the neural integrator is not tuned perfectly, but it is tuned well enough to extend the roughly 20-ms time constant of a typical neuron by a factor of about 1,000.

The fact that recurrent network models reproduce some of the core properties observed in biological integrator circuits has launched development of more detailed and realistic network models and testing the predictions of such models experimentally. These efforts also highlight the challenges involved in forging detailed links between the structure and function of neural circuits. Key questions remain even after decades of intensive study using a variety of systems and approaches.

For example, oculomotor integrator circuits typically contain two opposing classes of neurons, one increasing and the other decreasing their firing rates as eye position changes in a given direction. This arrangement is not restricted to oculomotor integrators but is also found in cortical regions implicated in decision making and working memory. Models have shown that mutual inhibition between these opposing populations can play a role in sustaining activity and integration. Although anatomical studies provide some support for this idea, studies in the goldfish showed that integration remains intact even when connections between the opposing populations are removed.

Another key question regards the mechanisms for tuning integrator networks. Experimental studies suggest that integrator networks are subject to modification via experience; in other words, they are tuneable. Although such tuning presumably occurs via changes in the strength of synaptic connections between neurons, direct evidence for this has yet to be obtained. In short, although much has been learned about how integration *could* be implemented, the details of the network architecture that actually support integration in any particular instance remain to be definitively established.

A detailed understanding of how we maintain the position of our eyes is an important end unto itself, with

clinical relevance. However, as pointed out earlier, the solutions found here may apply equally to cognitive functions including short-term memory and decision making. Optical imaging of large populations of neurons along with temporally precise manipulations of their activity and detailed anatomical reconstructions, combined with theoretical models of network function, may soon provide the answers.

Learning and Memory Depend on Synaptic Plasticity

Experience can modify neural circuits to support memory and learning (Chapter 3). It is generally believed that experience-dependent changes responsible for learning and memory occur primarily at synapses. Multiple forms of synaptic plasticity have been identified, and each of these presumably supports a different set of functions.

Just as there are multiple forms of plasticity, there are multiple forms of learning. Different forms of learning can be defined based on the amount and type of information provided. In supervised learning, explicit instruction is given about the behavior needed to perform a task. In reinforcement learning, on the other hand, only a positive reward or a negative punishment is provided to indicate whether that task is being performed properly. Finally, unsupervised learning involves no instructive information at all, but rather organizes input data on the basis of its intrinsic structure without supervision. In the following sections, we discuss an example of unsupervised learning involving Hebbian plasticity and an example of reinforcement learning in the cerebellum. (The various types of learning and memory and their cellular and circuit mechanisms are described in detail in Chapters 52–54.)

Dominant Patterns of Synaptic Input Can be Identified by Hebbian Plasticity

Cortical neurons receive synaptic input from thousands of other neurons and combine this information in patterns of action potentials. The strength of synaptic transmission at each of the synapses determines how the information arriving from many inputs is combined to affect the firing of the neuron. Setting the strength of all the synapses to zero would obviously make for a noninformative neuron of no functional use. Similarly, setting them to nonzero values that extract a signal dominated by random noise would also not produce a signal of value. Instead, neurons can best serve a useful function by extracting the most

interesting aspects of the information carried by their inputs. Theoretical analysis of a form of plasticity known as Hebbian indicates one way that this could happen in an unsupervised manner.

In 1949, Donald Hebb proposed that synapses should strengthen when a given presynaptic input to a neuron cooperates with a sufficient number of coactive inputs to cause that neuron to fire an action potential. Evidence for Hebbian synaptic plasticity has been obtained from many studies (Chapter 54). By itself, Hebbian plasticity would keep making synapses stronger and stronger, so some other form of plasticity must exist to prevent this from happening. Such compensatory forms of plasticity are called homeostatic, and experiments have revealed these forms of plasticity as well. Theoretical analysis indicates that a combination of Hebbian and homeostatic plasticity can adjust synapses, without any additional supervisory signal, so that they extract the combination of a neuron's inputs that is most highly modulated relative to other combinations (Figure 5-7). This is a reasonable candidate for the most interesting signal carried by those inputs, and thus, Hebbian plasticity provides a way for neurons to determine and extract such signals.

Synaptic Plasticity in the Cerebellum Plays a Key Role in Motor Learning

Although a detailed understanding of how the cerebellum contributes to complex human motor skills is lacking, a great deal is known about its role in simple forms of motor learning. Among the most thoroughly studied is a paradigm known as *delay eyeblink conditioning*, in which a neutral sensory stimulus such as a

light or a tone is repeatedly paired with an aversive unconditioned stimulus (US) such as an air puff to the eye. After several days of such training, animals learn to close their eye in response to the previously neutral stimulus (the light or tone), known as the conditioned stimulus (CS), in anticipation of the US (the air puff). The timing of the eyelid closure is highly specific to the delay between the onset of the CS and the US.

Eyelid conditioning has been an extremely useful paradigm for understanding cerebellar function because it maps onto the structure of cerebellar circuitry in a particularly clear way (Figure 5-8). Information about the CS is first encoded by cerebellar granule cells and then relayed to Purkinje cells. The US is encoded by a completely separate input pathway, known as the olivocerebellar or climbing fiber system. In contrast to the many thousands of inputs from granule cells, each Purkinje cell receives a single powerful climbing fiber input from a brain stem nucleus known as the inferior olive. Electrophysiological recordings revealed that climbing fiber inputs to one particular region of the cerebellum signal the occurrence of the US, that is, a stimulus that is irritating to the cornea. This discovery was made possible by the fact that the climbing fiber evokes a distinct suprathreshold response in the Purkinje cell known as a complex spike.

A key to understanding how the cerebellum mediates learning was the discovery that the complex spike triggers plasticity at synapses between granule cells and Purkinje cells. Specifically, the cooccurrence of input from a presynaptic granule cell and a complex spike in the postsynaptic Purkinje cell results in a persistent weakening of the granule cell input, a form of plasticity known as cerebellar

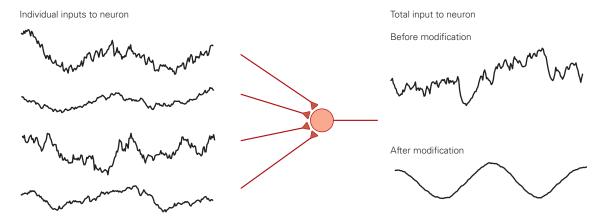


Figure 5–7 Hebbian plasticity can identify relevant input signals to a neuron. In this example, a neuron receives 100 inputs; firing rates for four of them are shown (*left*). Each of the input rates is noisy but contains, within the noise, a sinusoidal signal. The input rates are multiplied by synaptic strengths

(brown triangles) and then summed to produce the total input to the neuron (*right*). Before Hebbian plasticity occurs, the synapses have random weights, resulting in the noisy trace; after modification, the total input reveals the underlying sinusoidal signal.