Chapter 2

A Silicon Model of Auditory-Nerve Response

Nonlinear signal processing is an integral part of sensory transduction in the nervous system. Sensory inputs are analog, continuous-time signals with a large dynamic range, whereas central neurons encode information with limited dynamic range and temporal specificity, using fixed-width, fixed-height pulses. Sensory transduction uses nonlinear signal processing to reduce real-world input to a neural representation, with a minimal loss of information.

An excellent example of nonlinear processing in sensory transduction occurs in the cochlea, the organ that converts the sound energy present at the eardrum into the first neural representation of the auditory system, the auditory nerve. Humans can process sound input over a 120-dB dynamic range, yet the firing rate of an auditory-nerve fiber can encode only about 25 dB of sound intensity. Humans can sense binaural time differences of the order of 10 μs, yet an auditory-nerve fiber can fire at most once per millisecond. Using limited neural resources, the cochlea creates a representation that preserves the information essential for sound localization and understanding. Moreover, this neural code expresses auditory information in a way that facilitates feature extraction by higher neural structures.

This chapter describes the architecture and operation of an integrated circuit that models, to a limited degree, the evoked responses of the auditory nerve. The chip receives as input a time-varying voltage corresponding to sound input, and computes outputs that correspond to the responses of individual auditory-nerve fibers. The chip models the structure as well as the function of the cochlea; all subcircuits in the chip have anatomical correlates. The chip computes all outputs in real time, using analog continuous-time processing. The original research in

this chapter was done in collaboration with Carver Mead; parts of this chapter were originally published in (Lazzaro and Mead, 1989b).

2.1 Neural Architecture of the Cochlea

Both mechanical and electrical processing occur in biological cochleas. The sound energy present at the eardrum is coupled into a mechanical traveling-wave structure, the basilar membrane, which converts time-domain information into spatially-encoded information by spreading out signals in space according to their time scale (or frequency). Over much of its length, the velocity of propagation along the basilar membrane decreases exponentially with distance. The structure also contains active electromechanical elements; outer hair cells have motile properties, acting to reduce the damping of the passive basilar membrane and thus allowing weaker signals to be heard. Axons from higher brain centers innervate the outer hair cells; these centers may dynamically vary the local damping of the cochlea, providing frequency-specific automatic gain control (Kim, 1984).

Inner hair cells occur at regular intervals along the basilar membrane. Each inner hair cell acts as an electromechanical transducer, converting basilar-membrane vibration into a graded electrical signal. Several signal-processing operations occur during transduction. Inner hair cells half-wave rectify the mechanical signal, responding to motion in only one direction. Inner hair cells primarily respond to the velocity of basilar-membrane motion, implicitly computing the time derivative of basilar-membrane displacement (Dallos, 1985). Inner hair cells also compress the mechanical signal nonlinearly, reducing a large range of input sound intensities to a manageable excursion of signal level.

Spiral-ganglion neurons connect to each inner hair cell, and produce fixed-width, fixed-height pulses in response to inner-hair-cell electrical activity. The

synaptic connection between the inner hair cell and the spiral-ganglion neuron might implement a stage of automatic gain control, exploiting the dynamics of synaptic-transmitter release (Geisler and Greenberg, 1986). Auditory-nerve fibers are axons from spiral-ganglion neurons; these fibers present a neural representation of audition to the brain.

When pure tones are presented as stimuli, an auditory-nerve fiber is most sensitive to tones of a specific frequency. This characteristic frequency corresponds to maximum basilar-membrane velocity at the location of the inner hair cell associated with the nerve fiber. The spiral trunk of the auditory nerve preserves this ordering; the nerve fibers are mapped cochleotopically and tonotopically. The mean firing rate of an auditory fiber encodes sound intensity, over about 25 dB of dynamic range. The temporal pattern of nerve firings reflects the shape of the filtered and rectified sound waveform; this phase locking does not diminish at high intensity levels (Evans, 1982).

2.2 Silicon Models of the Cochlea

Both mechanical and electrical processing occur in the cochlea. In the chip, however, we model both types of computation using electronic processing. Lyon and Mead have designed a silicon model of the mechanical processing of the cochlea (Lyon and Mead, 1988a; Mead, 1989). Their circuit is a one-dimensional physical model of the traveling-wave structure formed by the basilar membrane. In this model of cochlear function, the exponentially tapered stiffness of the basilar membrane and the motility of the outer hair cells combine to produce a pseudoresonant structure.

A circuit model implements this view of cochlear hydrodynamics, using a cascade of second-order sections, with exponentially scaled time constants. This analog, continuous-time circuit model computes the pressure of selected discrete

points along the basilar membrane in real time. The cascade structure enforces unidirectionality, so a discretization in space does not introduce reflections that could cause instability in an active model. An amplifier in each second-order section provides active positive feedback in the circuit, modeling the active mechanical feedback provided by the outer hair cells in physiological cochleas. An automatic-gain-control system can increase sensitivity to weak sounds by locally varying the amount of positive feedback in the second-order sections. Appendix 2A provides a brief description of the circuits in the model.

This model of cochlear mechanics is the foundation of our model of auditory nerve response. We have not implemented an automatic-gain-control system to vary the damping of the structure locally; however, the chip has inputs that provide local control of membrane damping, allowing off-chip experiments with automatic gain control. To complete the circuit model of the auditory periphery, we have added circuits that model the functions of the inner hair cells and spiral-ganglion neurons; Figure 2.1 shows the complete architecture of the chip.

Our inner-hair-cell circuit models the signal-processing operations that occur during transduction: velocity-sensing, nonlinear compression, and half-wave rectification. Our spiral-ganglion-neuron circuit converts the analog output of the inner-hair-cell circuit into fixed-width, fixed-height pulses; the mean firing rate of the circuit encodes the intensity of inner-hair-cell circuit response, whereas the temporal pattern of pulses reflects the waveform shape of inner-hair-cell circuit response. Appendix 2B provides a description of these two circuits.

The synaptic connection between the inner hair cell and the spiral-ganglion neuron might implement a stage of automatic gain control, exploiting the dynamics of synaptic-transmitter release (Geisler and Greenberg, 1986). The nonlinear compression of our inner-hair-cell circuit approximates the static effect of the automatic gain control of the inner-hair-cell synapse. A direct
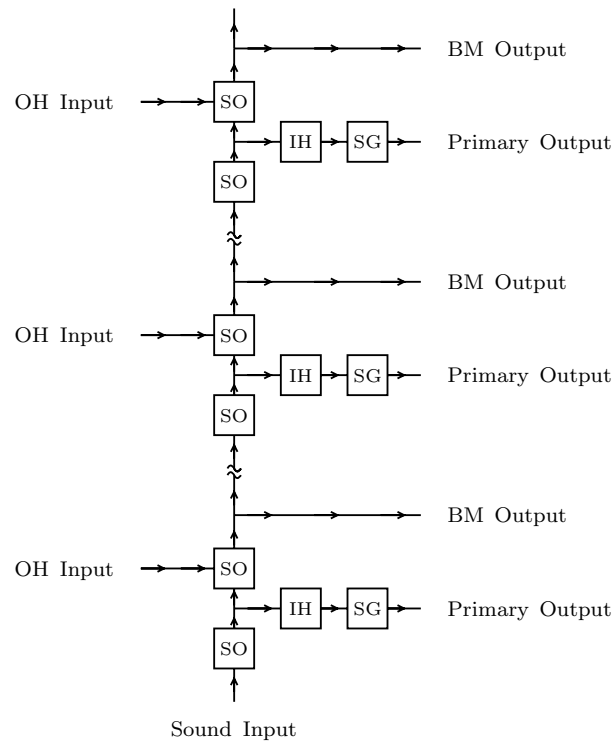
Figure 2.1. Block diagram of the auditory-nerve chip. Sound input travels down the basilar-membrane model, a cascade of second-order (SO) sections with exponentially increasing time constants. Basilar-membrane (BM) circuit outputs show pressure along the membrane, whereas inputs modeling efferent innervation of outer hair (OH) cells control local damping of the membrane circuit. Taps along the basilar membrane connect to a circuit model of inner hair (IH) cells; outputs from inner hair cells connect to circuits that model spiral-ganglion (SG) neurons. These neurons form the primary output of the chip, thus modeling auditory-fiber response.

implementation of dynamic automatic gain control would accurately model time adaptation and two-tone synchrony suppression; this enhancement is currently under development.

The outputs from spiral-ganglion-neuron circuits are the primary outputs of the chip. Additional outputs of the circuit display basilar-membrane pressure, for diagnostic purposes. The circuit we report in this chapter has eight auditory-fiber outputs; the projects described in Chapters 4 and 5 use cochlea circuits with 62 auditory-fiber outputs.

2.3 Silicon Basilar-Membrane Response

To test the tuning properties of the silicon auditory-nerve fibers, we duplicated a variety of classical auditory-nerve measurements. In these experiments, we tuned the basilar-membrane circuit to span about seven octaves, from 50 Hz to 10,000 Hz. We set the maximum firing rates of the auditory-fiber outputs at 150 to 300 spikes/s, with spike widths of 5 to 20 μs.

In this configuration, without an input signal, the auditory-fiber outputs fire at less than 0.1 spike/s. At the characteristic frequency of a fiber, pure tones of a few millivolts peak amplitude produce responses significantly above this spontaneous rate. The chip can process tones up to about 1 V of peak amplitude, yielding approximately 60 dB of usable dynamic range. Adding a preprocessor to the basilar-membrane circuit to limit intense input signals would extend the upper limit of the dynamic range. A biological cochlea has a mechanical limiter as a preprocessor — the stapedial reflex. Designing more sensitive inner-hair-cell circuits would extend the lower limit of dynamic range. Both dynamic-range enhancements are currently under development.

Figure 2.2(a) shows the amplitude of the silicon basilar-membrane response to pure tones, at a position with a best frequency of about 1900 Hz. We

set 0 dB at 3 mV peak, an input amplitude sufficient to produce responses above the spontaneous rate in auditory-fiber outputs near this basilar-membrane position. The 0-dB frequency-response curve shows a flat response for frequencies significantly below the best frequency, a 12-dB response peak at the best frequency, and a sharp dropoff to the noise floor for frequencies significantly above the best frequency. The 0-dB chip frequency-response curve is qualitatively similar to the 70-dB frequency-response curve taken from the basilar membrane of the squirrel monkey using the Mossbauer effect, shown in Figure 2.2(b) (Rhode 1971). Near the best frequency, basilar-membrane pressure, computed by the chip, is approximately equal to the basilar-membrane displacement, as measured by Rhode. Quantitatively, the bandwidth of the response peak of the chip is wider than that of the physiological data; a cascade of second-order sections does not yield an optimal model of cochlear hydrodynamics (Lyon and Mead, 1988b).

The 10-dB and 20-dB chip frequency-response curves show a decrease in the magnitude of the resonance peak, and a shift downward in best frequency. This qualitative behavior matches the nonlinear behavior of the squirrel-monkey basilar membrane. In the chip, the saturation of amplifiers that model the motility of the outer hair cells causes nonlinear behavior. A similar phenomenon may occur in the physiological system; for large sound intensities, outer hair cells may not be capable of a linear response to basilar-membrane motion.

Another explanation of nonlinear basilar-membrane response also is plausible. The efferent fibers that innervate the outer hair cells might be controlling frequency-selective automatic gain control. In this model, efferent fibers allow the outer hair cells to reduce the damping of the basilar membrane for soft sounds, whereas for louder sounds, efferent fibers inhibit the motility of outer hair cells, decreasing the size of the pseudoresonance peak (Kim, 1984). Our chip
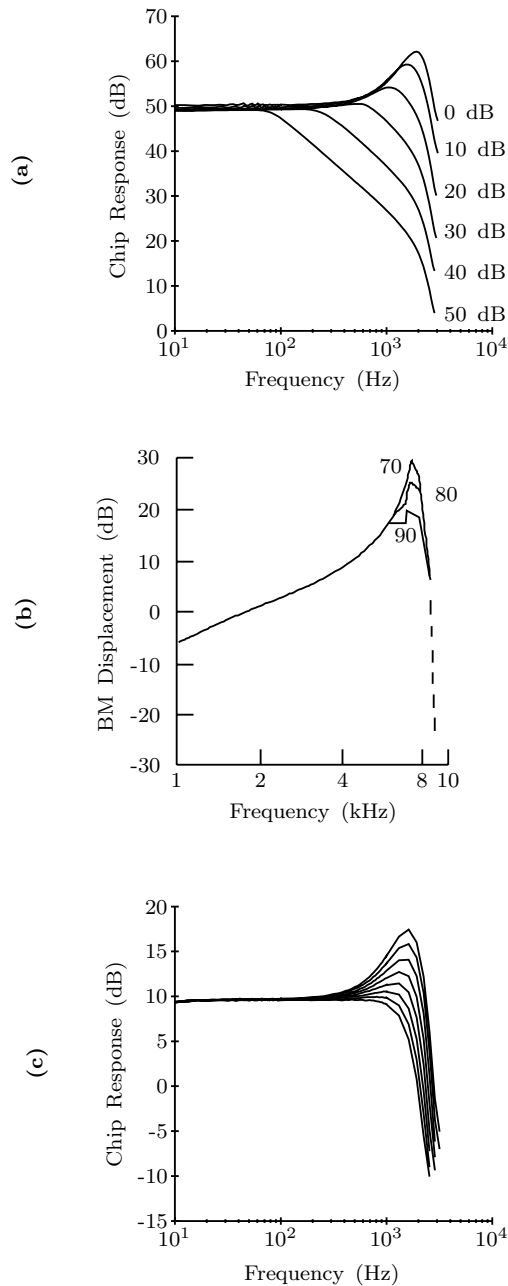
Figure 2.2. **a.** Plots showing the response of the basilar-membrane circuit at a single point, to pure tones at a fixed input amplitude (0 dB = 3 mV peak). The 0-dB to 40-dB curves are vertically shifted to match the low-frequency response of the 50-dB plot. **b.** Transfer functions of a single position on the basilar membrane of the squirrel monkey (Rhode, 1971). The curves show amplitude of vibration for constant malleus displacement. Two curves are vertically shifted to match the low-frequency response of the third curve. **c.** Plots showing the response of the basilar-membrane circuit to 10-dB pure tones, for different amounts of basilar-membrane damping at the measurement point.

does not model this automatic-gain-control system. There are, however, inputs into the chip that model the outer-hair-cell efferent fibers, allowing local control of damping at several points along the basilar-membrane model.

Figure 2.2(c) shows the effect of locally varying the strength of outer-hair-cell damping, near the position of recording, on the frequency response of the basilar-membrane model. The magnitude of the resonance peak decreases and the best frequency of the response shifts downward as the damping is increased. In the automatic-gain-control model, damping is increased for increasing signal levels; in this way, the model qualitatively matches the physiological responses of Figure 2.2(b). An integrated automatic-gain-control system for this chip is currently under development.

The frequency-response plots of the basilar-membrane model for 30-dB to 50-dB inputs, as shown in Figure 2.2(a), do not match physiological behavior. The saturation of the amplifiers that model the stiffness of the basilar membrane causes these undesired nonlinear effects. In practice, in many of the classical auditory-nerve experiments, the saturating nonlinearity of the inner-hair-cell circuit masks these nonphysiological effects. As mentioned earlier, future designs will incorporate more sensitive inner hair cells and a model of the stapedial reflex, to provide a large dynamic range without operating the basilar-membrane circuit in this regime; to manage this increased operating range, these designs will incorporate inner hair cells with dynamic automatic gain control.

## 2.4 Tuning Properties of Silicon Auditory-Nerve Fibers

We have characterized the tuning properties of the auditory-nerve-fiber circuit model, using both pure tones and clicks. In response to a click of medium intensity, a silicon auditory-nerve fiber produces one or several spikes. To extract the click response from these spikes, we present the click stimulus to the chip

many times, and record the responses of a silicon auditory-nerve fiber. These data are reduced to a poststimulus-time (PST) histogram, in which the height of each bin of the histogram indicates the number of spikes occurring within a particular time interval after the presentation of the click.

A PST histogram of the response of a silicon auditory-nerve fiber to a repetitive rarefaction click stimulus shows a half-wave-rectified version of a damped sinusoidal oscillation (Figure 2.3(a)). The frequency of this oscillation, 1724 Hz, is approximately the best frequency of the basilar-membrane position associated with this silicon nerve fiber. The half-wave rectification of the inner-hair-cell circuit removes the negative polarity of oscillatory waveform from the PST histogram of the click response. Repeating this experiment using a condensation click recovers the negative polarity of oscillation; a compound PST histogram, shown in Figure 2.3(b), combines data from both experiments to recreate the ringing waveform produced by the basilar-membrane circuit. Figure 2.3(c) shows a compound PST histogram of the click response of an auditory fiber in the cat (Kiang *et al.*, 1965). Qualitatively, the circuit response matches the physiological response.

Figures 2.3(a) and 2.3(b) are chip responses to a 60-mV click stimulus (26 dB, 0 dB = 3 mV peak). Higher-intensity clicks produce oscillatory responses with increased damping; a compound PST histogram of chip auditory-nerve response to a 36-dB click shows reduced ringing (Figure 2.3(d)). This effect is a direct result of the nonlinear response of the basilar-membrane model; physiological basilar-membrane click responses also show reduced ringing at high click-intensity levels (Robles *et al.*, 1976).

In response to a pure tone of sufficient intensity and appropriate frequency, the silicon auditory fiber produces spikes at a constant mean rate, as shown in Figure 2.4. The mean spike rate of a silicon fiber, in response to a constant tone,
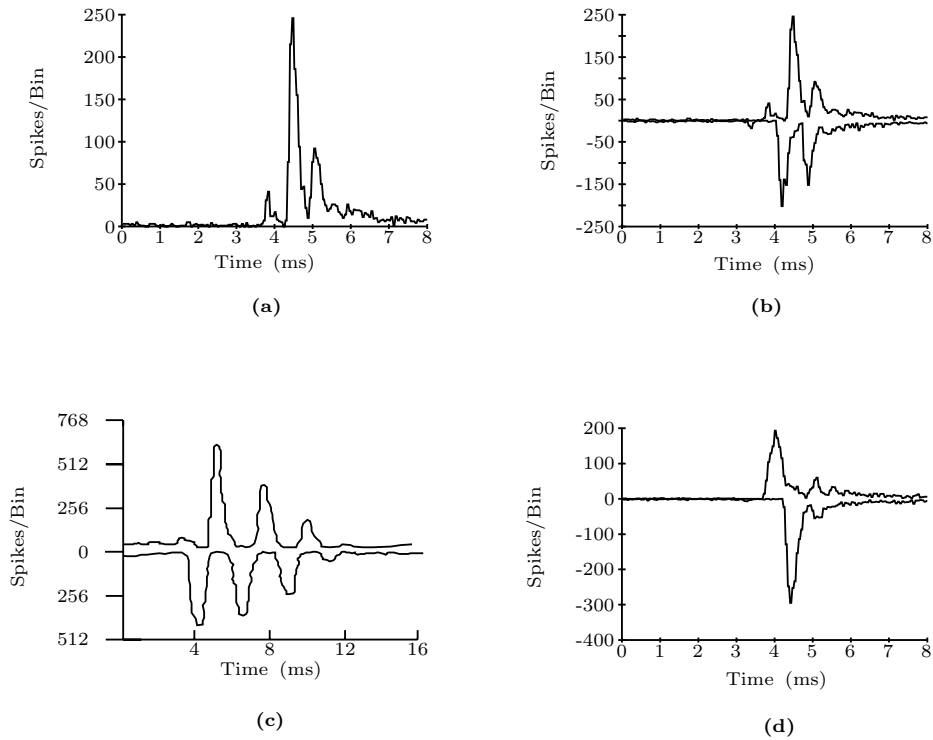
Figure 2.3. a. Post-stimulus time (PST) histogram of the rarefaction click response of a silicon auditory-nerve fiber. Click amplitude is 60 mV (26 dB peak); click width is 100 μs. Histogram is for 2000 click presentations; the width of each bin is 58 μs. **b.** Compound PST histogram of the click response of a silicon auditory-nerve fiber. Rarefaction click response is plotted as positive values; condensation click response is plotted as negative values. Conditions are identical to those of part a. **c.** Compound PST histogram of the click response of an auditory fiber in the cat (Kiang et al., 1965). Click level is 30 dB relative to threshold response level; click width is 100 μs. Rarefaction click response is plotted as positive values; condensation click response is plotted as negative values. **d.** Compound PST histogram of the click response of a silicon auditory-nerve fiber, for a 200-mV click (36-dB click). All other conditions are identical to those of part a.

does not decrease over time, unlike that of a physiological auditory fiber; this lack of adaptation indicates the absence of dynamic automatic gain control in our model.

Figure 2.5(a) shows the mean spike rate of a silicon auditory fiber as a function of pure tone frequency. For low-amplitude tones, the fiber responds to a narrow range of frequencies; for higher-intensity tones, the fiber responds to a wider range of frequencies. The saturating nonlinearities of the basilar-membrane circuit and of the inner-hair-cell circuit cause the bandwidth of the fiber to increase with sound intensity. Qualitatively, this behavior matches the iso-intensity plots from an auditory-nerve fiber in the squirrel monkey (Rose *et al.*, 1971), shown in Figure 2.5(b).

Figure 2.6(a) shows the mean spike rate of a silicon auditory fiber as a function of pure tone amplitude, at frequencies below, at, and above the best frequency of the fiber. In response to its characteristic frequency, 2100 Hz, the fiber encodes about 25 dB of tone amplitude before saturation. Figure 2.6(b) shows rate-intensity curves from an auditory fiber in the cat (Sachs and Abbas, 1974). At its characteristic frequency, the physiological fiber also encodes about 25 dB of tone amplitude before saturation. The shape of the biological and silicon curves at the characteristic frequency is remarkably similar, giving us some confidence in the validity of this modeling paradigm. In response to frequencies below and above the characteristic frequency, the functional forms of the silicon fiber responses are different from those of the physiological data. Most notably, the saturation rate of a silicon fiber for frequencies below the fiber's characteristic frequency exceeds the saturation rate of the silicon fiber at the fiber's characteristic frequency. This behavior is a direct result of the undesired saturation, at high input intensities, of amplifiers that model the stiffness of the
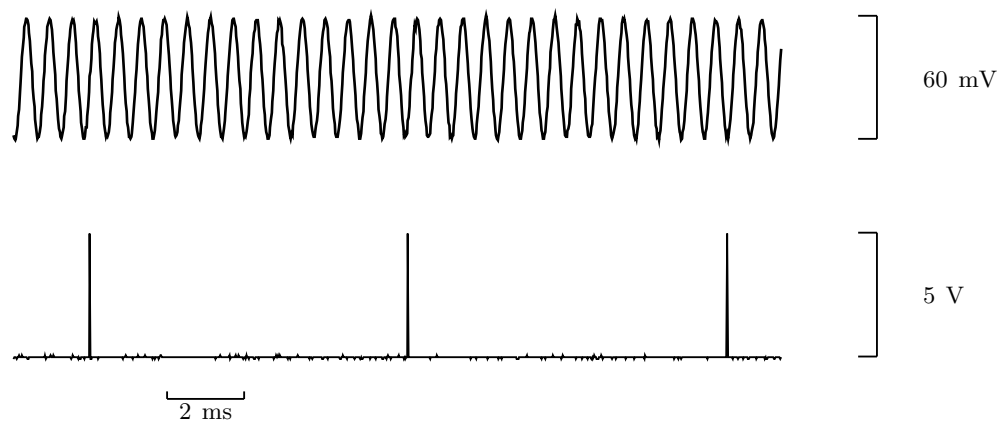
Figure 2.4. Output of a silicon auditory fiber (bottom trace) in response to a sinusoidal input (top trace). The frequency of the input is the characteristic frequency of the fiber.
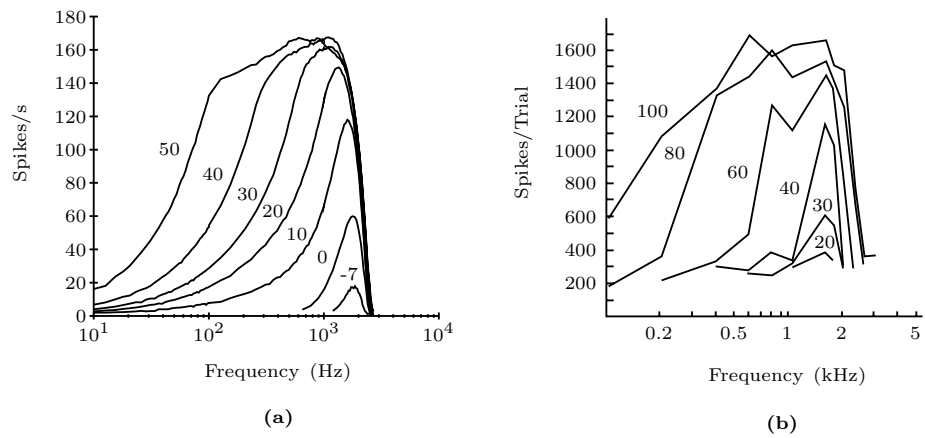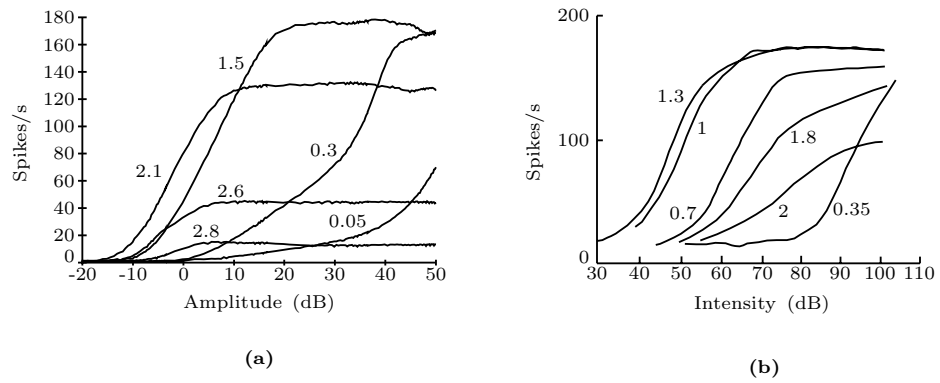
Figure 2.5. a. Plots showing the mean spike rate of a silicon auditory fiber as a function of pure tone frequency. Legend numbers indicate tone amplitude, in dB. **b.** Plots showing the number of discharges of an auditory fiber in the squirrel monkey, in response to a 10-s pure tone (Rose *et al.*, 1971). Legend numbers indicate tone amplitude, in dB.

basilar membrane. Above the best frequency of the silicon fiber, the response of the model decreases in a manner that is reminiscent of its biological counterpart.

Figure 2.7(a) shows iso-response curves for four silicon auditory-nerve fibers. These plots represent an iso-rate section through the iso-intensity curves of Figure 2.5(a), at a spike rate for each fiber that was comfortably above the spontaneous rate. The chip response accurately models the steep high-frequency tail of tuning curves from cat auditory fibers (Kiang, 1980), shown in Figure 2.7(b); the shapes of physiological and chip tuning curves are qualitatively similar. The bandwidth of the chip fibers for low sound intensities, however, is significantly wider than that of the physiological response. This problem stems from the wider bandwidth of the basilar-membrane circuit model, relative to that of the physiological data, as well as from the lack of a dynamic automatic-gain-control system for modulating the damping of the basilar-membrane circuit. The high-frequency cutoff of the iso-response curves, shown in Figure 2.7(a), is much steeper than is the cutoff of the iso-input curves shown in Figure 2.5(a). In a linear system, these two measurements would give identical results. The difference reflects the presence of a saturating nonlinearity in the system; the inner-hair-cell circuit and the basilar-membrane circuit provide this saturation function.

2.5 Timing Properties of Silicon Auditory-Nerve Fibers

As the click response of Figure 2.3 shows, the temporal firing patterns of the silicon auditory-nerve fibers encode information. Figure 2.8(a) shows period histograms of a chip fiber, in response to -5-dB to 50-dB pure tones at the fiber's characteristic frequency; these histograms show the probability of a spike output occurring within a particular time interval during a single cycle of the input sinusoid. The fiber preserves the shape of the input sinusoid throughout

Figure 2.6. a. Plots showing the mean spike rate of a silicon auditory fiber as a function of pure tone amplitude. Label numbers indicate tone frequency, in Hz. **b.** Plots showing the mean spike rate of an auditory fiber in the cat, as a function of pure tone amplitude (Sachs and Abbas, 1974). Label numbers indicate tone frequency, in Hz.
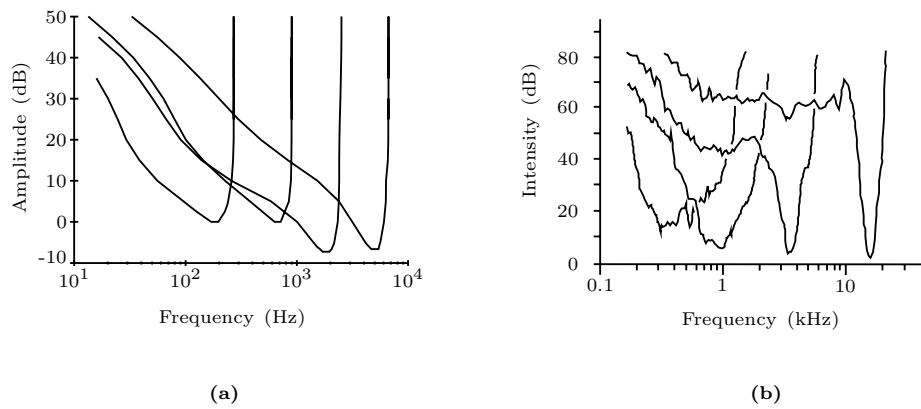
Figure 2.7. a. Plots showing iso-response curves for four silicon auditory fibers. The plots represent an iso-rate section through the iso-intensity curves of each fiber. Constant rates for each curve are, from the highest-frequency curve downward, 21.5, 16, 61, 59 spikes/s. **b.** Plots showing tuning curves from auditory fibers in the cat (Kiang, 1980). Fifty-ms tone bursts were presented at 10/s. Each tuning curve shows the sound pressure level (SPL) at the tympanic membrane (eardrum) that generates 10 spikes/s more activity during the tone bursts than during the silent interval.

this intensity range; this behavior matches data from an auditory fiber in the cat (Rose *et al.*, 1971), shown in Figure 2.8(b). Unlike the cat fiber, however, the silicon fiber does not preserve absolute phase at higher intensities; this deficiency results from the saturation of the second-order section amplifiers that model basilar-membrane stiffness.

The temporal firing patterns of the silicon auditory-nerve fiber are, however, a good representation of signal periodicity; the synchronization ratios (normalized magnitude of the first Fourier coefficient) of the period histograms in Figure 2.8(a) are 0.5 to 0.6, comparable to those of physiological data at the same frequency. The firing patterns of silicon nerve fibers maintain partial synchrony in the presence of masking noise. Figure 2.9(a) shows the degree of synchronization of a silicon fiber to a pure tone at the best frequency of the fiber, in the presence of white noise. Figure 2.9(b) shows similar data from the auditory nerve of the squirrel monkey (Rhode *et al.*, 1978). Qualitatively, the data from the chip fiber and those from the cat fiber are similar.

The temporal firing patterns of the silicon auditory-nerve fiber preserve the shape of complex waveforms; Figure 2.10(a) shows the compound period histograms of a silicon fiber, in response to two harmonically related tones, combined with equal amplitudes. A Fourier analysis of the histogram shows strong peaks at 868 Hz and at 1156 Hz, along with distortion products; the fiber preserves periodicity information. Figure 2.10(b) shows an auditory fiber in the cat, which preserves periodicity information of a tone pair in the same manner (Goblick and Pfeiffer, 1969).
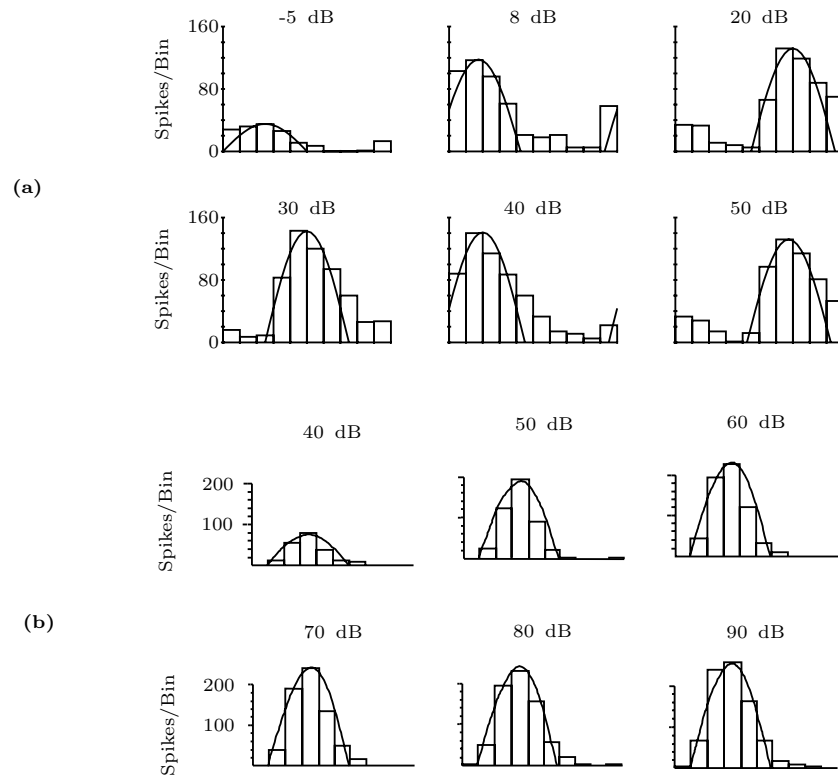
Figure 2.8. a. Period histograms of the silicon auditory-fiber response to a pure tone of 1840 Hz, near the fiber's best frequency. Amplitude of tone is shown above each plot. Histogram width is 54 μs. Each histogram begins at a constant position, relative to the input sinusoid; each is fitted to a sinusoid of best amplitude and phase. **b.** Period histograms of the response of an auditory fiber in the cat, to a low-frequency tone (Rose *et al.*, 1971). Amplitude of pure tone is shown above each plot. Each histogram is fitted to a sinusoid of best amplitude but fixed phase.
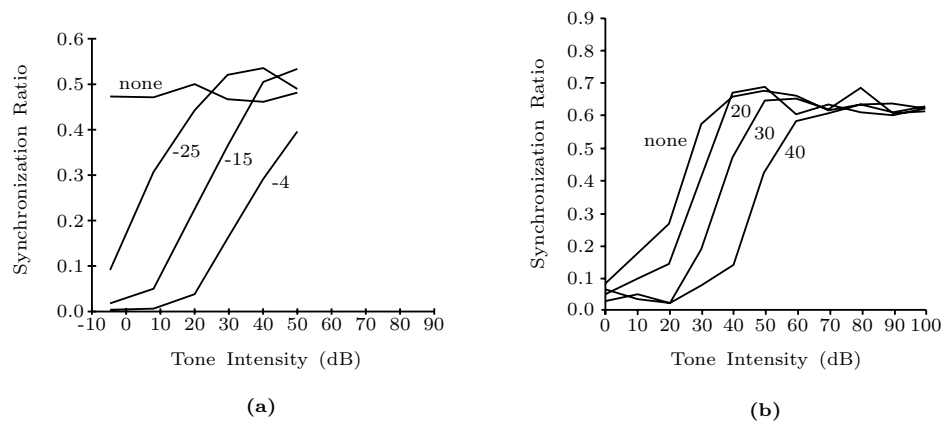
Figure 2.9. a. Plots of the synchronization ratio of PST histograms of a silicon auditory-fiber response to pure tones. Pure tone frequency is 1840 Hz, near the fiber's best frequency. Synchronization ratio is plotted as a function of tone amplitude. White noise is mixed with the tone; the legend indicates the amplitude of the noise signal, in $dB/\sqrt{Hz}$. **b.** Plots of the synchronization ratio of an auditory fiber in the squirrel monkey (Rhode *et al.*, 1978). Pure tone frequency is 900 Hz, the fiber's best frequency. Synchronization ratio is plotted as a function of tone amplitude. Band-limited (100- to 2000-Hz) noise is mixed with the tone; the legend indicates the spectrum density of the noise signal, in dB.

(a)   (b)
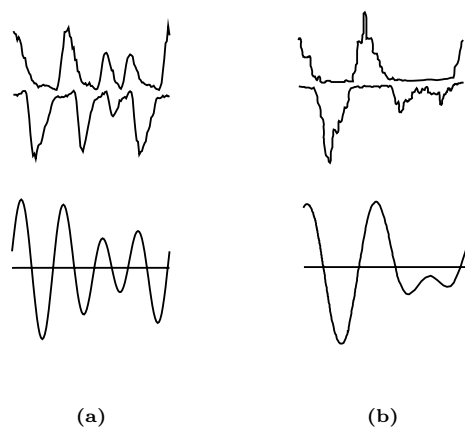
Figure 2.10. a. Top plot shows the compound period histogram of silicon auditory nerve response to two tones, at 868 Hz and 1156 Hz ($f_3 : f_4$), mixed at equal amplitudes (24 dB). Bottom plot shows the best-fit combination of 868 Hz and 1156 Hz to the histogram (arbitrary amplitude and phase). The best frequency of the fiber is about 1900 Hz. **b.** Top plot shows the compound period histogram of a cat auditory-nerve-fiber response to two tones ($f_2 : f_3$), at 538 Hz and 807 Hz (Goblick and Pfeiffer, 1969). Bottom plot shows the best-fit combination of 538 Hz and 807 Hz to the histogram (arbitrary amplitude, but the same relative phase as the stimulus).

2.6 Discussion for the Scientist

The integrated circuit captures many essential features of data representation in the auditory nerve; moreover, it computes the representation in real time. The performance of the present circuit, however, falls short of the performance of the auditory nerve in several significant respects. Each shortcoming in performance can be traced to the incomplete modeling of an aspect of cochlear physiology.

The bandwidth of the resonant peak in the basilar-membrane circuit is wider than that of physiological basilar-membrane response; a cascade of second-order sections does not yield an optimal model of cochlear hydrodynamics (Lyon and Mead, 1988b). The bandwidth of the silicon auditory-nerve-fiber tuning curves is also wider than that of physiological fiber response. The insufficient basilar-membrane circuit bandwidth obviously is a factor in this shortcoming; the lack of a dynamic automatic-gain-control system for modulating the damping of the basilar-membrane circuit is another important factor.

This absence of this automatic-gain-control system, as well as the absence of circuits that model the dynamics of the synaptic connection between the inner hair cell and the spiral-ganglion neuron, result in other performance deficiencies. Specifically, a silicon auditory nerve fiber does not exhibit time adaptation, and does not exhibit physiological two-tone rate supression and two-tone synchrony supression.

Finally, a silicon auditory fiber lacks the wide dynamic range of physiological cochleas; the chip can process sounds that range over only 60 dB. Improved inner-hair-cell circuits that are sensitive to smaller voltage excursions would improve circuit dynamic range, as would a circuit model of the stapedial reflex. With a stapedial-reflex circuit as a preprocessor, the undesired saturation of amplifiers that model the stiffness of the basilar membrane would not occur; this saturation causes the unwanted phase shift of the period histograms of

Figure 2.8(b), which show the temporal firing patterns of a silicon auditory nerve in response to sinusoids at varying intensities. This saturation also causes the saturation rate of a silicon fiber in response to frequencies below the fiber's characteristic frequency to exceed the saturation rate of a silicon fiber in response to the fiber's characteristic frequency, as shown in Figure 2.6(b).

Future research on this project involves implementing these enhancements to our model of auditory-nerve response. Modeling neural systems, in a physical medium that shares many of the strengths and weaknesses of the biological substrate, offers a unique perspective on the relationship of neural structure and function. For example, during the design of our silicon auditory nerve model, many iterations of the inner-hair-cell circuit design were required to achieve acceptable circuit performance. The major difficulty in circuit design was modeling the long-term adaptation of the cell to static cilia deflection. Duplicating this property of the cell was crucial to the successful operation of the chip. Without this auto-zeroing property, the inner-hair-cell circuit cannot adjust for variances in circuit elements due to fabrication tolerances; as a result, most auditory-fiber circuit outputs either refuse to fire or fire at a maximum rate, regardless of sound input.

Sensing basilar-membrane velocity, rather than pressure, provides this auto-zeroing function; our design goal evolved to be the synthesis of a sensitive time-differentiator circuit. Our inner-hair-cell circuit includes a time-differentiation circuit that works by comparing the instantaneous value of a voltage that represents basilar-membrane pressure with a time-averaged value of this voltage. Thus, the characteristics of the silicon medium dictated the functional form of our model. We believe that any medium imposes a direction on modeling; it is thus advantageous to choose a modeling medium — analog VLSI technology — that shares characteristics with the biological system under study.

2.7 Discussion for the Engineer

Our integrated-circuit model captures many essential features of data representation in the auditory nerve; moreover, it computes the representation in real time. There are many traditional engineering representations of audition, however, that are also amenable to analog implementation. What advantages does a silicon auditory-nerve representation offer to a designer of artificial sensory systems?

As shown in Figures 2.3, 2.8, and 2.10, an auditory-nerve fiber encodes a filtered, half-wave–rectified version of the input waveform, over a wide dynamic range, using the temporal patterning of fixed-width, fixed-height pulses. This representation supports the efficient, massively parallel computation of signal properties, using autocorrelations in time and cross-correlations between auditory fibers. In this representation, a correlation is simply a logical AND operation, performed by a few synapses in neural systems, or by a few transistors in silicon systems. Axonal delays in neural systems provide the time parameter for computing autocorrelations; in silicon systems, we model this delay with compact monostable circuits (Mead, 1989). We use these techniques in the projects in Chapter 4 and Chapter 5.

The nonlinear filtering properties of the auditory-nerve fibers, shown in Figures 2.5 and 2.7, enhance these correlations. In a quiet environment, auditory fibers have narrow bandwidths; each fiber carries independent information, yielding rich correlations. In noisier environments, the tuning of auditory fibers widens, increasing the number of fibers that carry information about the signal. This detuning ensures that some fibers still encode signal properties reliably (Greenberg, 1988).

As shown in Figure 2.6, auditory fibers encode about 25 dB of signal intensity. Dynamic automatic gain control, present in a physiological cochlea,

enhances this range; in addition, different populations of auditory fibers have different thresholds, further enhancing the encoding of signal intensity. Although not sufficient as a primary representation of sound, rate encoding of signal intensity is a valuable secondary cue, particularly for the detection of rapid spectral changes and the encoding of aperiodic sounds. Future versions of our chip will include these enhancements for rate encoding of signal intensity.

In conclusion, we have designed and tested an integrated circuit that computes, in real time, the evoked responses of auditory nerve, using analog, continuous-time processing. The chip offers a robust representation of audition, which can serve as a solid foundation for analog silicon systems that model higher auditory function.

Circuit Description of the Basilar-Membrane Model

The circuit model of cochlear mechanics developed by Lyon and Mead (Lyon and Mead, 1988a; Mead, 1989) is the foundation of our silicon model of auditory-fiber response. This appendix provides a brief description of the implementation of this basilar-membrane model.

The circuit is a one-dimensional physical model of the traveling-wave structure formed by the basilar membrane. In this viewpoint of cochlear function, the exponentially tapered stiffness of the basilar membrane and the motility of the outer hair cells combine to produce a pseudoresonant structure. The basilar-membrane circuit model implements this view of cochlear hydrodynamics using a cascade of second-order sections with exponentially scaled time constants.

Figure 2A.1 shows the CMOS circuit implementation of a second-order section. Input and output signals for the circuit are time-varying voltages. The gain blocks are transconductance amplifiers, operated in the subthreshold regime. Capacitors are formed using the gate capacitance of n-channel and p-channel MOS transistors in parallel. Because of subthreshold amplifier operation, the time constant of the second-order section is an exponential function of the voltage applied to the transconductance control inputs of $A1$ and $A2$, labeled τ in Figure 2A.1. Thus, a cascade of second-order circuits, with a linear gradient applied to the τ control inputs, has exponentially scaled time constants. To implement this gradient, we used a polysilicon wire that travels along the length of circuit, and that connects to the τ control input of each second-order section. A voltage difference across this wire, applied from off the chip, produces exponentially scaled time constants. The amplifier $A3$ provides active positive
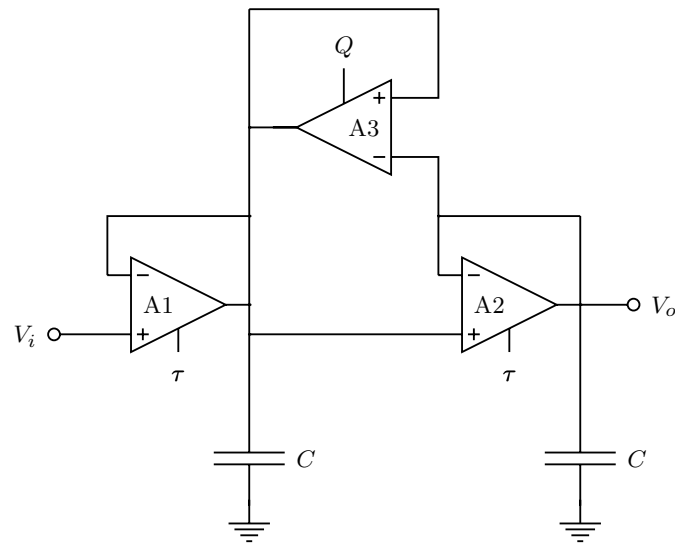
Figure 2A.1. Circuit implementation of a second-order section. Input V_i and output V_o are time-varying voltages. The τ and Q control inputs set bias currents on transconductance amplifiers $A1$, $A2$, and $A3$, to control both the characteristic frequency and the peak height of the lowpass-filter response.

feedback to the membrane, modeling the active mechanical feedback provided by the outer hair cells in biological cochleas. A second polysilicon wire is connected to the transconductance inputs of the $A3$ amplifiers in each second-order section (labeled Q in Figure 2A.1); a voltage gradient across this wire similar to that on the τ control inputs sets all the second-order sections to the same response shape.

A way to model the adjustment of basilar-membrane damping by higher brain centers is to use an automatic-gain-control system that varies the damping of the second-order sections locally. We have not implemented this automatic-gain-control system; however, we have brought off-chip several taps from the polysilicon wire that connects to the Q control of the second-order sections, allowing off-chip experiments with automatic gain control.

Appendix 2B

Circuit Description of the Auditory-Transduction Model

Our inner-hair-cell circuit models the signal-processing operations that occur during transduction: velocity sensing, nonlinear compression, and half-wave rectification. Our spiral-ganglion-neuron circuit converts the analog output of the inner-hair-cell circuit into fixed-width, fixed-height pulses; the mean firing rate of the circuit encodes the intensity of inner-hair-cell circuit response, whereas the temporal pattern of pulses reflects the waveform shape of inner-hair-cell circuit response. This appendix provides a description of the implementations of the inner-hair-cell circuit and of the spiral-ganglion-neuron circuit.

Figure 2B.1 shows our inner-hair-cell circuit model. A hysteretic-differentiator circuit (Mead, 1989) processes the input-voltage waveform from the basilar-membrane circuit, performing time differentiation and logarithmic compression. The circuit enhances the zero-crossings of the input waveform, accentuating phase information in the signal. The output voltage of the hysteretic differentiator connects to a novel implementation of a half-wave current rectifier.

Figure 2B.2 shows our half-wave current-rectifier circuit. To understand its operation, we consider the state of this circuit when the input voltage V_h is constant. If V_h is constant, $I_h = 0$, and V_a adapts such that $I_p = I_n$. For $I_h = 0$, we define the quiescent conditions $I_q \equiv I_p = I_n$ and $V_q \equiv V_a$. The value of I_q depends on the circuit bias voltage, V_s. A current mirror reflects this quiescent current to the circuit output. Thus, the output of the half-wave current-rectifier circuit in response to a constant voltage input is an adjustable bias current.

Now consider the circuit state when the input voltage V_h is a time-varying waveform. During the positive-going phase of the waveform, the current I_h is
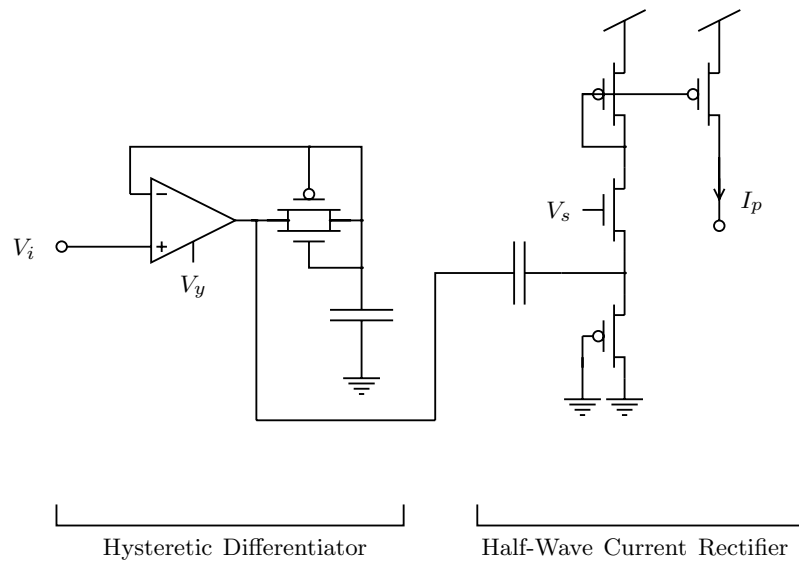
Figure 2B.1. The inner-hair-cell circuit model. Input V_i, from the basilar-membrane circuit, is a time-varying voltage. The hysteretic-differentiator circuit, biased by voltage V_y, performs time differentiation and logarithmic compression. The output of the hysteretic differentiator, a time-varying voltage, connects to the half-wave current-rectifier circuit, which is shown in more detail in Figure 2B.2.

positive, and $I_n = I_h + I_p$. As I_n increases, V_a must also increase; the amount of increase depends on the circuit bias voltage, V_s, as shown in the bottom graph in Figure 2B.2. If V_a increases, however, then I_p must decrease. So, during the positive-going phase of the waveform, the output current I_p decreases from the quiescent current I_q.

During the negative-going phase of the waveform, the current I_h is negative, $I_p = |I_h| + I_n$, and the output current of the circuit increases from the quiescent current I_q. Thus, the circuit converts the input time-varying voltage waveform V_h into a unidirectional current waveform I_p. For large $|I_h|$ relative to I_q, the current waveform I_p is not symmetrical about I_q, and the average value of I_p is greater than that of I_q; thus, the circuit performs the rectification function, as shown in the top graph in Figure 2B.2.

The current I_p is the output of the inner-hair-cell circuit. The spiral-ganglion neuron circuit model, shown in Figure 2B.3, converts this current into fixed-width, fixed-height pulses. The circuit — a slightly modified version of the neuron circuit described in (Mead, 1989) — creates a pulse rate that is linear in input current, for sufficiently low pulse rates. Thus, the average pulse rate of the circuit reflects the average value of I_p, whereas the temporal placement of each pulse reflects the shape of the current waveform I_p.
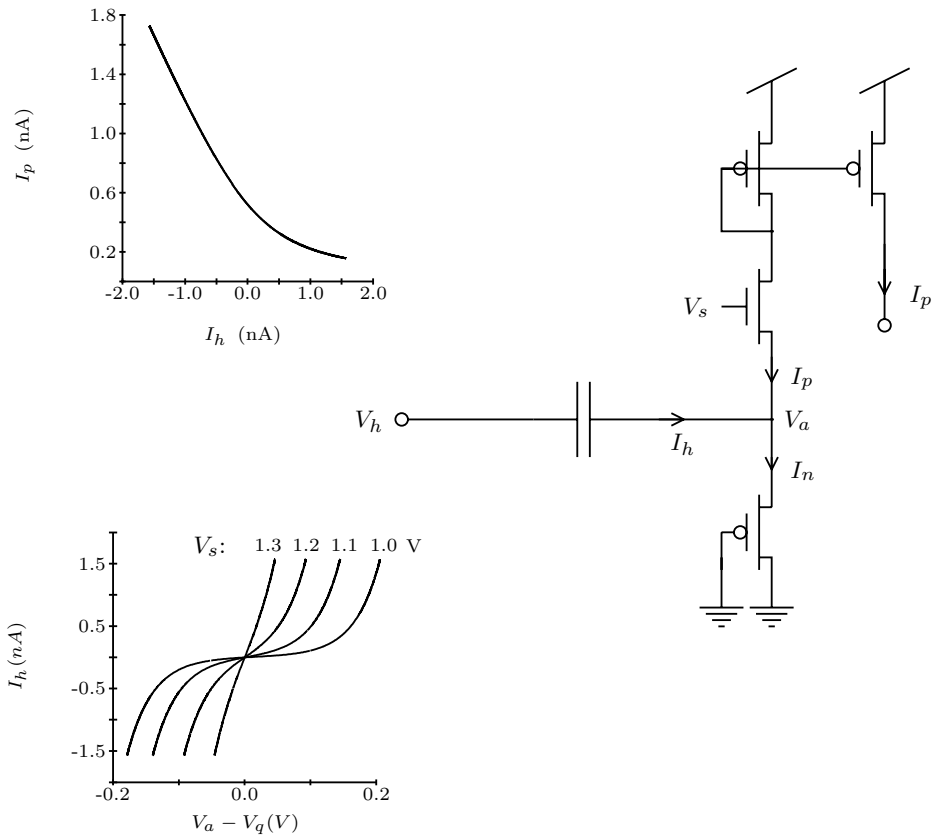
Figure 2B.2. The half-wave current-rectifier circuit. Input V_h, from the hysteretic-differentiator circuit, is a time-varying voltage. A floating capacitor couples V_h into the node associated with V_a, as the bidirectional time-varying current I_h. The bottom graph shows the change in V_a required to sink or source I_h, for several values of bias voltage V_s; the voltage V_q is the value of V_a when $I_h = 0$. When $V_a = V_q$ and $I_h = 0$, the circuit output, the unidirectional current I_p, is at a quiescent value, I_q, set by V_s. Nonzero values of I_h modulate the output current I_p about I_q; for large $|I_h|$ relative to I_q, the circuit output I_p is a half-wave–rectified version of I_h, as shown in the top graph. Graphs show theoretical responses.
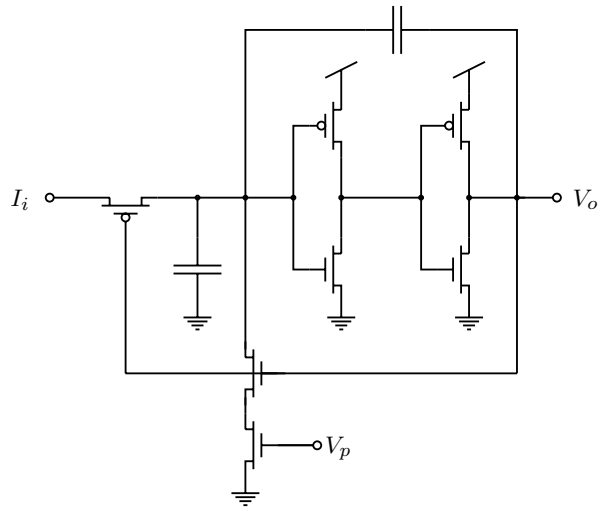
Figure 2B.3. The spiral-ganglion-neuron circuit. Circuit input, from the half-wave–rectification circuit, is the unidirectional current I_i. The circuit converts this current into fixed-width, fixed-height voltage pulses, at output V_o. The bias voltage V_p sets pulse width; the output voltage V_o pulses between V_{dd} and ground.