

A Silicon Model of an Auditory Neural Representation of Spectral Shape

John Lazzaro¹

California Institute of Technology

Pasadena, California, USA

Abstract

The paper describes an analog integrated circuit that implements an auditory neural representation of spectral shape. The circuit contains silicon models of the cochlea, inner hair cells, spiral ganglion cells, and the neurons that compute an amplitude-invariant representation of spectral shape. The chip uses the temporal information in each silicon auditory-nerve fiber to compute this final representation. The chip was fabricated and fully tested; the paper includes data comparing the silicon auditory-nerve representation and the final representation. The 9000 transistor chip computes all outputs in real time using analog continuous-time processing.

1. Introduction

The cochlea is the sense organ of hearing. It converts acoustic signals into the first neural representation of audition; the auditory nerve, containing about 50,000 fibers in man, carries this representation to the brain. Outputs from the left and right cochleas serve as inputs for the neural structures that perform spatial sound localization and sound understanding. In addition, several species of animals use their cochleas as sensors for active sonar processing.

Sound recognition, sound localization, and active sonar are practical and interesting engineering endeavors. There is renewed interest by the engineering community in understanding biological approaches to these problems and in adapting these biological solutions to engineering systems [1]. The first task is difficult, because of the incomplete knowledge of the structure and function of auditory processing in the brain. The second task is also difficult, because of the large computational demands of neural processing.

We are exploring analog VLSI technology as a computational medium for auditory neural processing. Analog VLSI offers real-time, low-power computation of neural algorithms, and shares many of the computational properties of the biological substrate [2]. The effort began with silicon models of the hydrodynamics of the cochlea [3] and of auditory-nerve response [4]. We used these cochlear circuits as components in silicon models of auditory lateralization [5] and of pitch perception [6].

We believe that sound recognition, like visual object recognition, benefits from multiple representations of sensory input. These representations should be dedicated to the robust

¹ Present Address: Computer Science Division, EECS, University of California at Berkeley.

extraction of different properties of the input. As a general-purpose vision system should have separate representations for form, color, texture, motion, and depth, an auditory system should have separate representations for properties like amplitude modulation, frequency modulation, spatial location, periodicity pitch, and spectral shape.

Several amplitude-invariant representations of spectral shape have been proposed in the auditory literature [7]; these proposals include neural algorithms for computing the new representation from the auditory-nerve input. In this paper, we report on an analog integrated circuit that implements one of these algorithms. The circuit contains 9000 transistors, and computes the representation in real time, using analog, continuous-time processing. The circuit was fabricated and fully tested; we present data showing the performance of the device.

2. Representations for Spectral Shape

The auditory-nerve representation itself codes the spectral shape of input signals. Spectral signal processing in the cochlea begins before electrical transduction; sound is coupled into a 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 [8].

In signal processing terms, a point along the basilar membrane acts as a low-pass filter with a resonant peak and a sharp cutoff. The resonant frequency of the low-pass filter decreases exponentially at points progressively distant from the mechanical input. Inner hair cells, distributed at regular intervals along the basilar membrane, act as electromechanical transducers, converting basilar-membrane vibrations into graded electrical signals [9]. Synapses from spiral-ganglion neurons connect to the inner hair cells; most auditory-nerve fibers sending signals to the brain are axons from these spiral-ganglion neurons.

Unlike inner hair cells, the auditory-nerve signals are not graded electrical potentials; the auditory-nerve fibers produce fixed-width, fixed-height pulses in response to inner-hair-cell electrical activity. One measure of auditory-nerve tuning is the mean spike rate of an auditory nerve in response to sinusoids of different frequencies. Measured in this way, auditory-nerve fibers act as bandpass filters with gradual low-frequency cutoffs and sharp high-frequency cutoffs, reflecting the resonant peak and sharp cutoff of the basilar membrane. Different auditory-nerve fibers are tuned to different frequencies, associated with the position of its inner hair cell on the basilar membrane [10].

In this way, the auditory-nerve response represents the spectral shape of the input signal. However, this representation of spectral shape is not amplitude invariant. At higher amplitudes, the nerve fiber saturates, and the low-frequency cutoff of the filter response shifts grossly downward. At sound pressure levels (SPL) typical of normal conversation, about 60% of the auditory-nerve fibers are saturated. This fact is paradoxical, given that psychoacoustic experiments show that speech intelligibility improves with increased SPL [7].

Because of these characteristics, many auditory theorists consider a mean rate encoding of spectral shape insufficient for explaining auditory perception. Several theories involve the extraction of information from the fine-time structure of auditory-nerve outputs. Auditory-nerve fibers with resonant frequencies below 5kHz fire with much greater probability on one polarity of an input waveform. More specifically, the probability density function for spike generation is roughly a half-wave rectified version of the electrical waveform at the inner hair cell. This phase encoding of the signal persists in fully saturated fibers [10].

Several proposed representations for spectral shape involve comparing or combining the synchrony of firing between fibers connected to different inner hair cells [11,12,13]. Other proposed representations involve extracting information from the phase encoding of fibers connected to the same inner hair cell [14, 15, 16, 17, 18].

The simplest of the latter schemes involves connecting an auditory-nerve fiber with a resonant frequency of f_j to a matched filter for a spike repetition rate of $1/f_j$ [15]. A simple realization of this matched filter is a correlator that receives as input the auditory-nerve fiber response delayed by a time $1/f_j$, and the undelayed auditory-nerve fiber response [18]. The frequency characteristic of this autocorrelator shows strong peaks at frequencies nf_j , for positive integer n . The cochlear filter, however, removes all frequencies substantially above f_j from the input, and the frequency response of the system shows a single peak at f_j .

Auditory neurons rarely have saturated firing rates above 300 spikes per second; if firing in response to sinusoids of higher frequencies, spikes are synchronized to the input, but do not occur on every cycle of the input waveform. Combining the outputs of several auditory-nerve fibers connected to the same inner hair cell increases the likelihood of nerve firings on successive cycles of a high-frequency input sinusoid, allowing the matched filter to function correctly.

A criticism of matched filter is neurophysiological implausibility; in specific, how does a neuron that implements a matched filter know the proper f_j [7]? One can imagine an adaptive neuron, that learns the correct f_j by experience. Alternatively, one can imagine time delays hard-coded into the neural circuit. How plausible could the latter method be, given the inherent offsets and drift of neural circuitry? Subthreshold analog VLSI shares these problems with the neural substrate, and is an ideal medium to test the robustness of an algorithm to component tolerances. In addition, the matched filter does not require large numbers of wires to communicate between channels, and is a good engineering fit for VLSI, where wiring costs dominate component costs. We have designed, fabricated, and tested an analog VLSI chip that implements this algorithm.

3. Chip Architecture

Figure 1 shows a block diagram of the chip. The electrical signal representing sound input connects to a silicon model of the mechanical processing of the cochlea [3]. 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; in Figure 1, each box marked with an arrow represents a second-order section. The cascade structure enforces unidirectionality, so a discretization in space does not introduce reflections that could cause instability in an active model. This analog, continuous-time circuit model computes the pressure at selected discrete points along the basilar membrane in real time. There are 32 second-order sections on the chip.

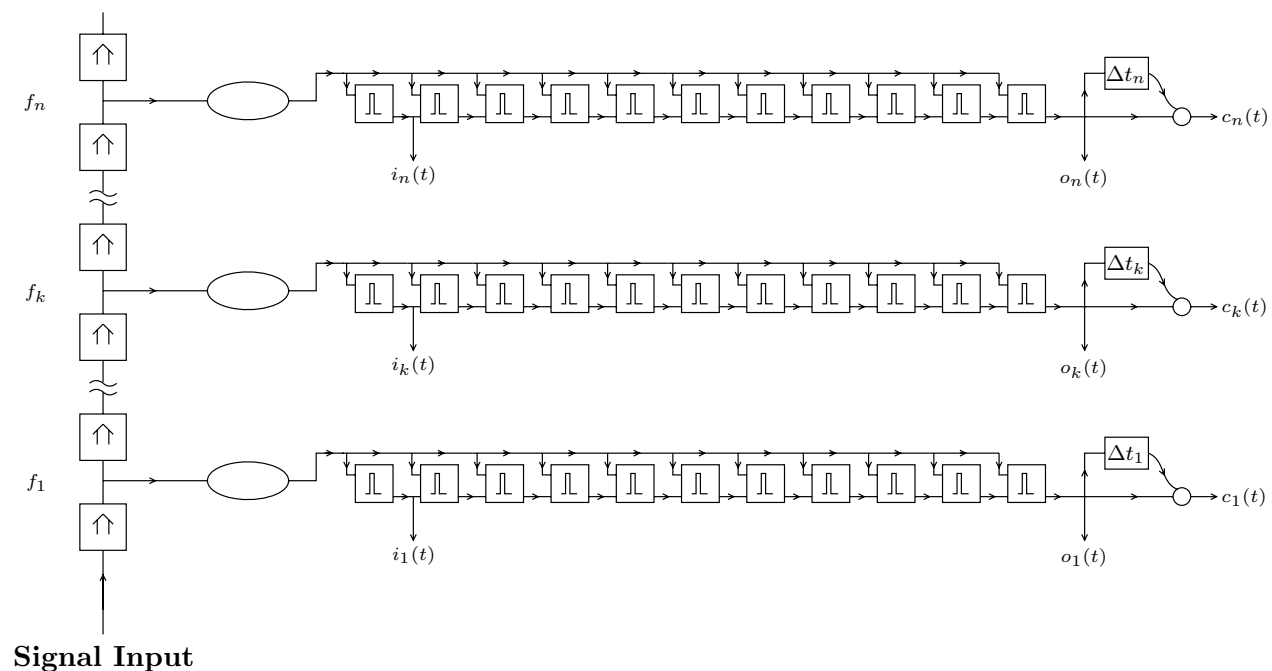


Figure 1. Block diagram for the chip. Sounds input enters a silicon cochlea [3], drawn as a cascade of second-order sections; each square box marked with an arrow represents a second-order section. The symbol f_j denotes the center frequency at positions along the silicon cochlea. Outputs from the silicon cochlea connect to circuit models of inner-hair-cell transduction [4], drawn as an ellipse. This circuit connects to circuits modeling spike generation and combination, drawn as boxes marked with a pulse [20]. The signal $i_j(t)$ represent the activity of a single spiral-ganglion-neuron circuit; the signal $o_j(t)$ represent the combined response of 11 spiral-ganglion-neuron circuits. This combined response is correlated with a time-delayed version of the combined response, to yield the final output $c_j(t)$. Time-delay circuits, drawn as boxes marked by the symbols Δt_j , delay the combined response by a time inversely proportional to the center frequency f_j . The correlation circuit, drawn as a small circle, performs a Boolean AND operation on the delayed and undelayed signals.

The output of each second-order section connects to a circuit that models the signal-processing operations that occur during inner-hair-cell transduction [4]; each ellipse in Figure 1 represents a inner-hair-cell circuit. 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 [9]. Inner hair cells also compress the mechanical signal nonlinearly, reducing a large range of input sound intensities to a manageable excursion of signal level. Our inner-hair-cell circuit performs these operations.

The output of the inner-hair-cell circuit connects to circuits that model spiral-ganglion neurons; each box marked with a pulse represents a spiral-ganglion-neuron circuit. The spiral-ganglion-neuron circuit converts the graded output of inner-hair-cell circuit into fixed-width, fixed-height pulses. The matched-filter algorithm requires combining the pulses produced by the spiral-ganglion neurons connected to the same inner hair cell. The spiral-ganglion-neuron circuit combines an external pulse input with its internally generated pulse; the final output from the cascade of 11 spiral-ganglion-neuron circuits combines the pulses of all of the neurons.

The combined pulsed output is delayed by a time matched to the resonant frequency f_j of its associated cochlear tap; the box marked with the symbol Δt_j performs this delay. A correlation neuron, drawn as a circle, takes the delayed and undelayed combined inputs, and produces the final system output $c_j(t)$. The intermediate signals $i_j(t)$ and $o_j(t)$ are also brought off chip.

The size of the chip is $2220\mu \times 2250\mu$; the chip was fabricated in a 2μ CMOS n -well low-noise MOSIS process. There are 32 matched filters on the chip; six outputs are brought off the chip on separate pads. Three $i_j(t)$ outputs and one $o_j(t)$ output are also brought off the chip.

4. Circuit Implementation

Figure 2 and Figure 3 show the circuit implementations of all the building blocks in Figure 1. As all of the circuits have been published previously, the descriptions in this section are not lengthy; the references provide additional details.

Figure 2(a) shows the CMOS circuit implementation of a second-order section [3], with input V_i and output V_o . 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 node labeled τ . 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 connects to the τ control input of each second-order section. A voltage difference across this wire, applied from off chip, produces exponentially scaled time constants. The amplifier controlled by the voltage q provides active positive feedback, modeling the active mechanical feedback provided by the outer hair cells in biological cochleas. A second polysilicon wire is connected to the q control input in each second-order section; 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.

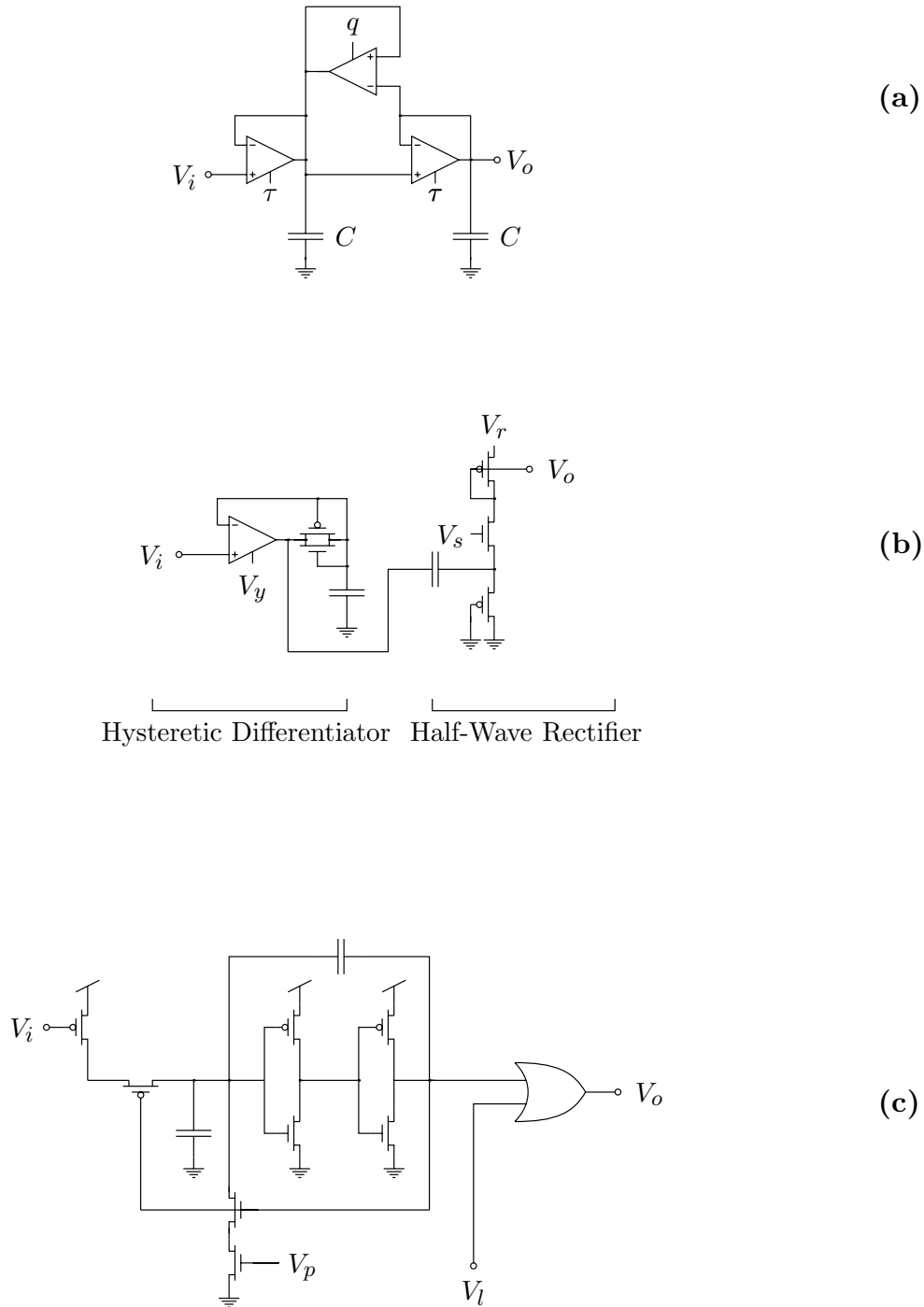


Figure 2. Circuit elements used in Figure 1. **(a)** Second-order section used in silicon cochleas, with voltage input V_i , voltage output V_o , and control voltages τ and q . Amplifiers are wide-range transconductance amplifiers [3]. **(b)** Circuit model of inner-hair-cell transduction, with voltage input V_i , voltage output V_o , and control voltages V_y , V_r , and V_s . **(c)** Spike generation and combination circuit, with voltage input V_i , previous pulse input V_l , and combined pulse output V_o . Control voltage V_p sets pulse width; OR gate is implemented with standard static CMOS logic circuits.

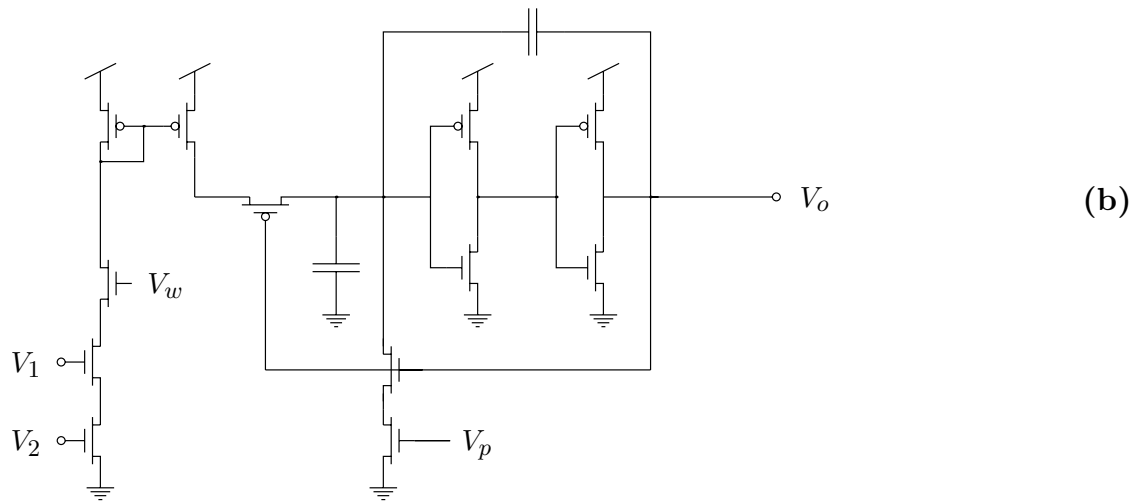
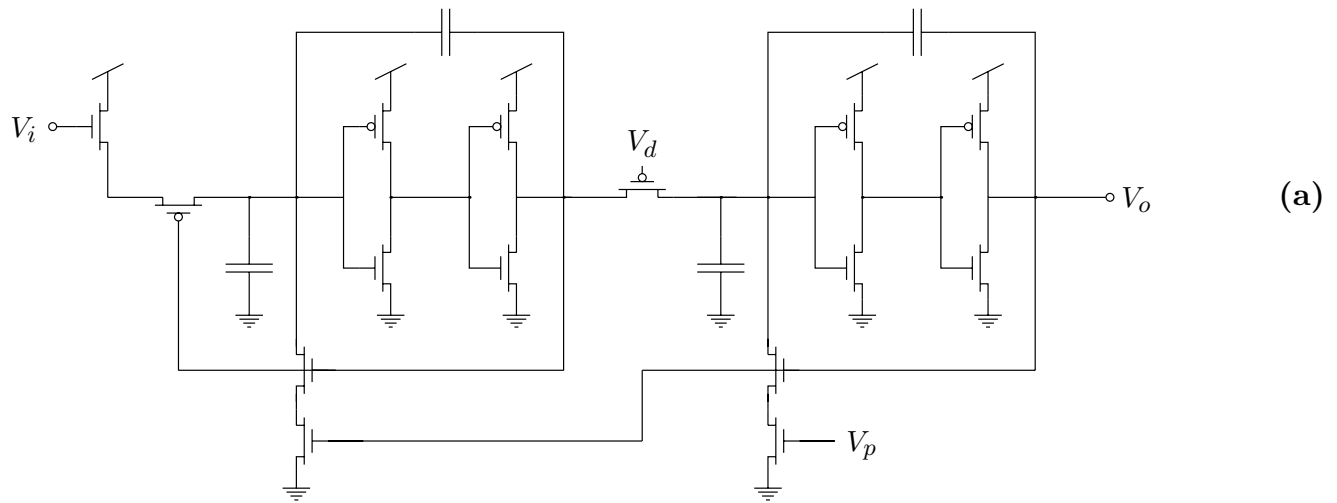


Figure 3. Circuit elements used in Figure 1. **(a)** Delay element, with pulse input V_i and pulse output V_o . Control voltage V_d sets pulse delay; control voltage V_p sets output pulse width. **(b)** Correlator circuit, with pulse inputs V_1 and V_2 , and pulse output V_o . Control voltage V_w scales output pulse rate; control voltage V_p sets output pulse width.

Figure 2(b) shows the inner-hair-cell model, with input V_i and output V_o . A hysteretic-differentiator circuit [19] processes V_i , performing time differentiation and logarithmic compression. The circuit enhances points in the waveform where the first derivative changes sign, accentuating phase information in the signal. The output voltage of the hysteretic differentiator connects to a half-wave current rectifier [4]. In this circuit, current from positive voltage transients is shunted to ground, while current from negative voltage transients passes through the p -channel transistor whose gate is labeled V_o . This transistor is best considered one-half of a simple current mirror; the other half of the current mirror is a part of the spiral-ganglion-neuron circuit.

Figure 2(c) shows the spiral-ganglion-neuron circuit, with input V_i and pulse output V_o . The circuit converts the input voltage into a unidirectional current, then converts this current into fixed-width, fixed-height pulses. The circuit — a slightly modified version of the neuron circuit in [20] — 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 input, whereas the temporal placement of each pulse reflects the shape of the input current waveform. A Boolean OR gate, implemented with standard static CMOS logic, combines these pulses with the output of the last spiral-ganglion neuron, presented on input V_l .

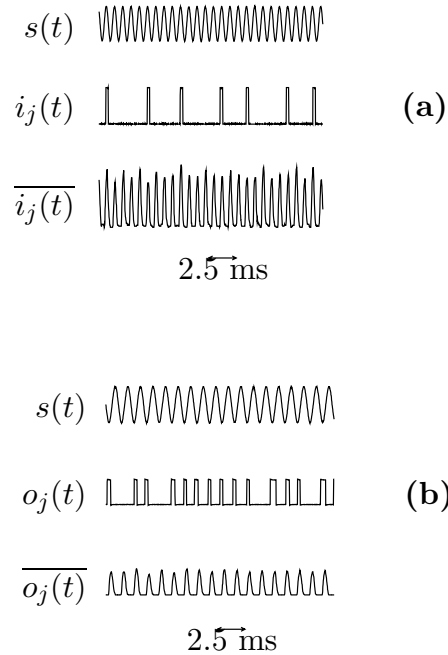


Figure 4. (a) Data from an $i_j(t)$ output of the chip, associated with a cochlear tap with $f_j = 1338\text{Hz}$, when driven with a sinusoid of frequency f_j and amplitude of 5mV peak. Trace $s(t)$ is chip input, trace $i_j(t)$ is output response, trace $\overline{i_j(t)}$ is the average output of $i_j(t)$, in response to many presentations of $s(t)$. (b) Data from an $o_j(t)$ output of the chip, associated with a cochlear tap with $f_j = 889\text{Hz}$, when driven with a sinusoid of frequency f_j and amplitude of 5mV peak. Trace $s(t)$ is chip input, trace $o_j(t)$ is output response, trace $\overline{o_j(t)}$ is the average output of $o_j(t)$, in response to many presentations of $s(t)$.

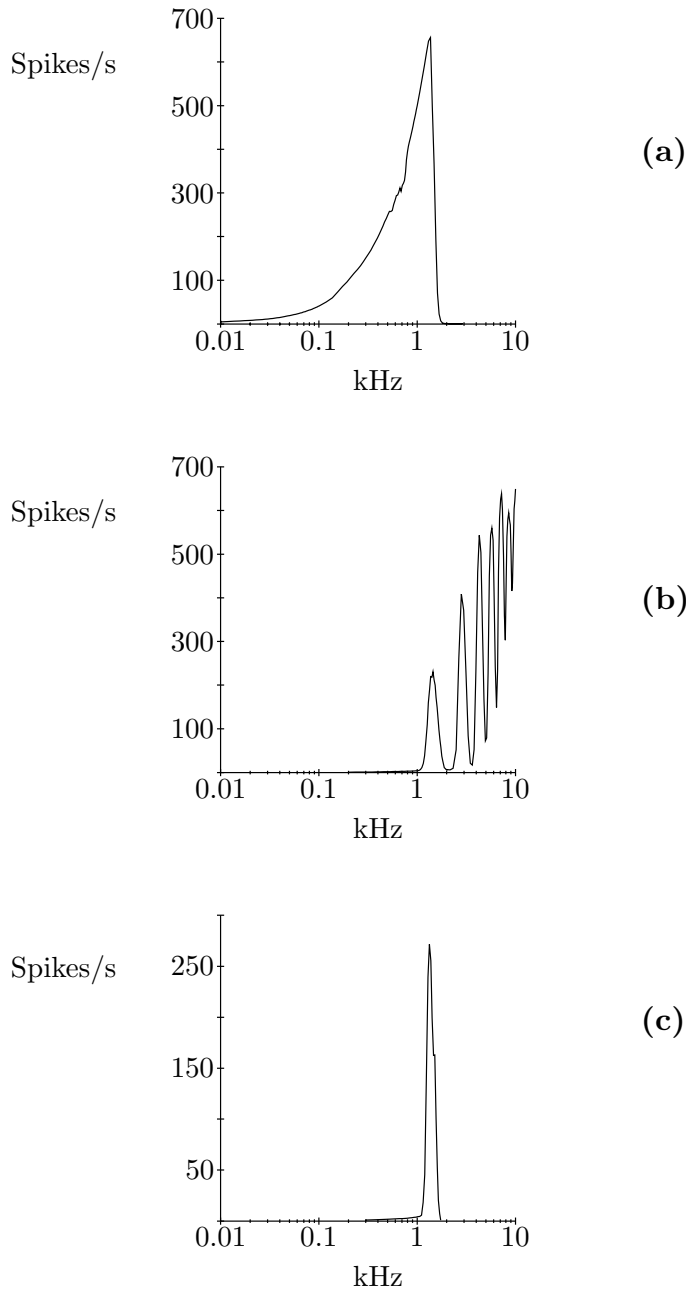


Figure 5. (a) Data from an $i_j(t)$ output of the chip, with $f_j = 1338\text{Hz}$. Plot shows average firing rate of output in response to a sinusoid at various frequencies, with amplitude of 10mV peak. (b) Data from a $c_j(t)$ output corresponding to the $i_j(t)$ output in (a), with the cochlear filtering action disabled. Experiment identical to (a). (c) Data from a $c_j(t)$ output corresponding to the $i_j(t)$ output in (a), with the cochlear filtering enabled. Experiment identical to (a).

Figure 3(a) shows the time-delay circuit for pulses, with pulse input V_i and pulse output V_o . The circuit is a single stage of the axon circuit shown in [20]. The control voltage V_d sets the delay time, while the control voltage V_p sets the output pulse width. The voltage V_d is set in the subthreshold range; as a result, the delay time is an exponential function of V_d . To match the delay time to the cochlear tuning, we connect the V_d node of all the delay circuits to a polysilicon wire, and apply a voltage gradient to this wire. This linear voltage gradient imposes an exponential gradient on the delay circuits, that can be adjusted to match the exponential scaling of the silicon cochlea. The transistor whose gate connects to control voltage V_d sets the time delay, and has a width of 6μ and a length of 12μ . These dimensions were chosen to improve the matching of the time-delay circuits without seriously increasing circuit layout area.

Figure 3(b) shows the correlator circuit, with inputs V_1 and V_2 and output V_o . The transistors associated with V_1 and V_2 implement the correlation, computing the Boolean AND function. These transistors produce a current, weighted by the control voltage V_w , if pulses are present on both V_1 and V_2 . This current is mirrored and presented to a neuron circuit, as in [20].

To produce uniform behavior for all $c_j(t)$ outputs, we applied voltage gradients along the cochlear dimension to several control inputs using polysilicon wires. The pulse width control voltage V_p for the time-delay circuits and the spiral-ganglion-neuron circuits were connected together on such a gradient line. The scaling control voltage V_w for the correlator circuit used a gradient line, as did the scaling control voltage V_r for the half-wave rectification circuit.

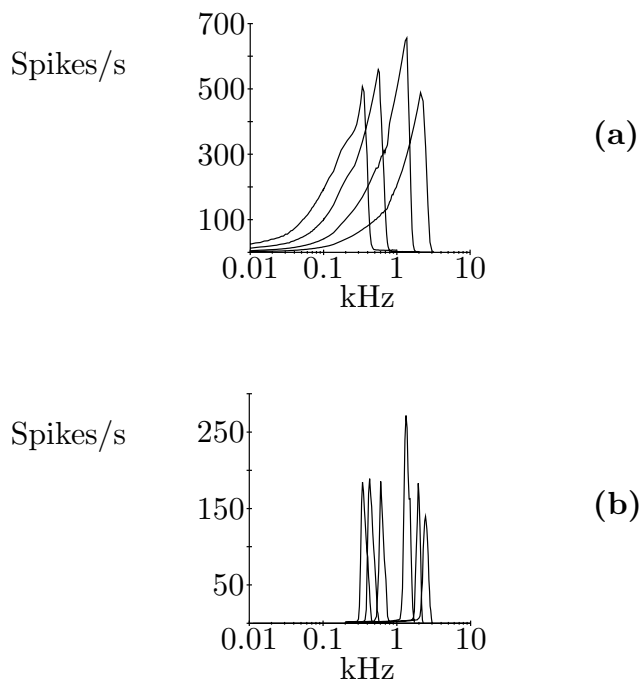


Figure 6. Data from several $i_j(t)$ and $c_j(t)$ outputs; control voltage settings are identical for all curves. **(a)** Data from four evenly spaced $i_j(t)$ outputs; plots shows average firing rate of outputs in response to a sinusoid at various frequencies, with amplitude of 10mV peak. **(b)** Data from six $c_j(t)$ outputs; first three curves and last three curves are evenly spaced along the cochlea. Experiment identical to (a).

5. Data

We tuned the silicon cochlea to span a decade in frequency, from approximately 400Hz to 4kHz. The high-frequency limit is approximately the highest frequency of phase locking in auditory-nerve fibers. The limited frequency range is required to insure correct resonant behavior in the small silicon cochlea (32 second-order sections). Maximum spike rates of the spiral-ganglion-neuron circuits averaged 550 spikes/second, slightly higher than biological spiral-ganglion neurons. We chose this rate so that the three distinct regions of algorithm operation would be represented: input signal frequency less than, equal to, and greater than saturated spike rate.

We set the gradient of the delay circuits to match the frequency gradient of the silicon cochlea. We set the pulse width of the spiral-ganglion-neuron circuits to $2\mu\text{s}$, much shorter than biological pulse widths ($300\mu\text{s}$). This short pulse width permits the use of a simple, two transistor correlator circuit (Figure 3(a)); a correlator circuit that operated on rising edges of pulses would function with pulse widths on the order of biological neurons.

Figure 4(a) shows a silicon auditory-nerve output, $i_j(t)$, if stimulated with a sinusoidal input $s(t)$. The frequency of $s(t)$ is the resonant frequency of the fiber, 1338Hz. Spikes do not occur every cycle, but spikes have a preferred position relative to the input. Averaging $i_j(t)$ over many presentations of $s(t)$ reveals this preference, as shown in $\overline{i_j(t)}$.

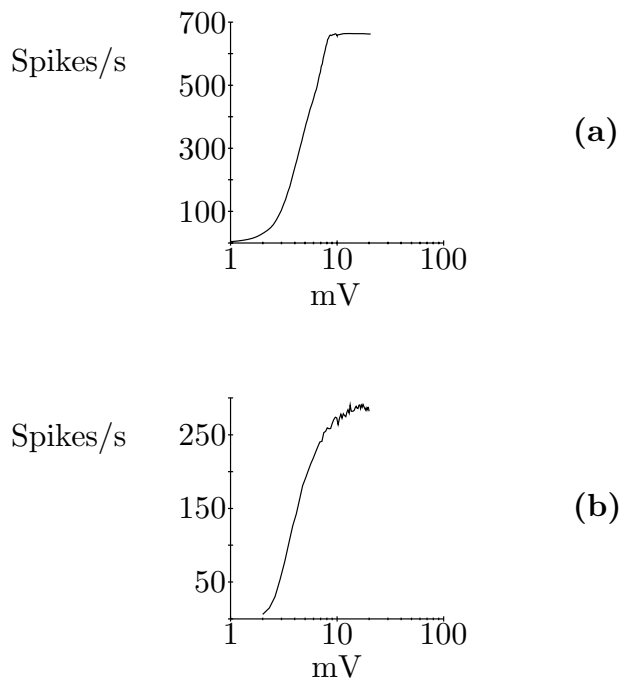


Figure 7. Data from the $i_j(t)$ and $c_j(t)$ outputs examined in Figure 5. **(a)** Plots show average firing rate of the $i_j(t)$ output in response to a sinusoid at various frequencies, with amplitudes of 2mV, 5mV, and 20mV peak; wider curves correspond to higher amplitudes. **(b)** Data from the $c_j(t)$ output; experiment identical to (a).

Figure 4(b) shows a combined pulse output, $o_j(t)$, if stimulated with a sinusoidal input $s(t)$ of frequency 889Hz, the resonant frequency of the fiber. Spikes occur nearly every cycle of the the input waveform, as needed for the correlation algorithm to function. The averaged waveform $\overline{o_j(t)}$ shows the synchrony between $s(t)$ and $o_j(t)$.

Figure 5(a) shows the frequency response of a silicon auditory-nerve fiber. We presented high amplitude (10mVpeak) sinusoids of different frequencies to the chip, and recorded the mean firing rate of the $i_j(t)$ output. Figure 5(b) shows the frequency response of the $c_j(t)$ output associated with $i_j(t)$, after disabling the frequency tuning of the silicon cochleas. The plot shows frequency peaks at $f_j, 2f_j, 3f_j, \dots$, as expected for an autocorrelator. Figure 5(c) shows the frequency response of the $c_j(t)$ output with the silicon cochlea properly tuned. The silicon cochlea removes the frequencies capable of exciting the $2f_j, 3f_j, \dots$ peaks of the autocorrelator, leaving a single sharp peak at f_j .

Figure 6(a) shows the frequency response of four $i_j(t)$ outputs that span the silicon cochlea. Figure 6(b) shows the frequency response of six $c_j(t)$ outputs that span the silicon cochlea. The $c_j(t)$ outputs all have suppressed $2f_j, 3f_j, \dots$ peaks and sharp f_j peaks, showing the good match between silicon cochlea tuning and autocorrelator delay times.

Figure 7(a) shows the frequency response of a silicon auditory-nerve fiber at three different input amplitude levels. At the lowest level (2mV peak), the range of frequencies that excites the fiber to at least 20% of its peak value span 763Hz. At the moderate level (5mV peak), this bandwidth is 1007Hz, and at the highest level (20mV peak), this bandwidth is 1372Hz. Measured in this way, the spectral selectivity of the output is sensitive to input amplitude.

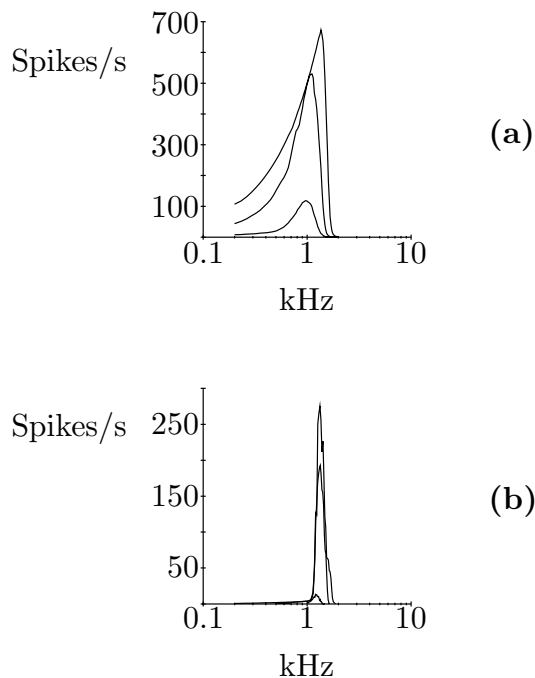


Figure 8. Data from the $i_j(t)$ and $c_j(t)$ outputs examined in Figure 5. **(a)** Plot shows average firing rate of the $i_j(t)$ output in response to a sinusoid of frequency f_j , at various peak amplitudes. **(b)** Data from $c_j(t)$ output; experiment identical to (a).

Figure 7(b) shows an experiment identical to Figure 7(a), performed on the correlator output associated with the auditory-nerve output of Figure 7(a). At the lowest level (2mV peak), the range of frequencies that excites the correlator output to at least 20% of its peak value span 356Hz. At the moderate level (5mV peak), this bandwidth is 432Hz, and at the highest level (20mV peak), this bandwidth is 457Hz. This bandwidth is relatively amplitude invariant, in comparison to the auditory-nerve measurements.

Figure 8(a) shows the amplitude response of the auditory-nerve output used in Figure 7(a). We presented sinusoids with different amplitudes to the chip, and recorded the mean firing rate of the output; the frequency of the sinusoid was f_j , the resonant frequency of the $i_j(t)$ output. The figure shows that the amplitude values used in Figure 7 correspond to the threshold of the response (2mV), the logarithmic region of the response (5mV), and the saturated region of the response (20mV). Figure 8(b) shows the amplitude response of the same correlator output as Figure 7(b). The correlator output preserves the sigmoidal response of the auditory-nerve output.

Figure 9 shows the effect of changing the pulse width of the spiral-ganglion-neuron circuits and delay circuits on the correlator output. Larger pulse widths produce larger bandwidths, as expected. The range of frequencies that excites the fiber to at least 20% of its peak value varies from 386Hz, for a pulse width of $0.5\mu\text{s}$, to 1028Hz, for a pulse width of $19\mu\text{s}$.

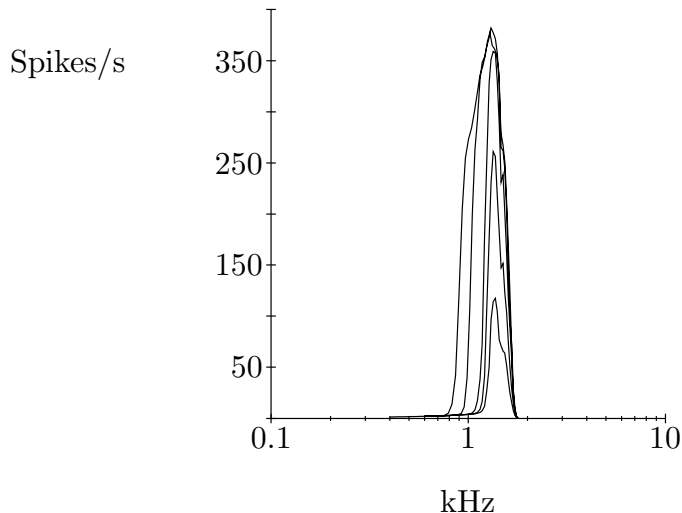


Figure 9. Data from the $c_j(t)$ output examined in Figure 5, showing the effect of changing spike pulse width on $c_j(t)$ response. Each plot corresponds to a different pulse width, and shows average firing rate of output in response to a sinusoid at various frequencies, with amplitude of 10mV peak. Widest curve corresponds to a pulse width of $19.1\mu\text{s}$; narrower curves correspond to pulse widths of $10.7\mu\text{s}$, $3.6\mu\text{s}$, $1.3\mu\text{s}$, and $0.5\mu\text{s}$.

6. Discussion

The circuit implements the matched-filter algorithm correctly and robustly. Figure 7(b) shows the desired amplitude-invariant frequency response, the major improvement over the auditory-nerve representation of Figure 7(a). Figure 6(b) shows the algorithm functions correctly, even if delays are implemented with a non-adaptive gradient that controls imprecise components. Figure 5(b) reveals the robust property of the algorithm; the first undesired peak $2f_j$ occurs a distance f_j away from the desired peak at f_j . The associated cochlear filter may have a cutoff frequency significantly greater than f_j , but significantly less than $2f_j$, without adversely affecting the correlator output $c_j(t)$.

The $c_j(t)$ outputs operate correctly over a limited range of input amplitudes (2 – –20mV), as shown in Figure 8. This property is not a limitation of the autocorrelation implementation, but of the implementation of the silicon cochlea. A second-order-section circuit with improved dynamic range and better saturation behavior would extend the range of the $c_j(t)$ outputs.

The circuit models known and proposed auditory function at a high level of abstraction. We chose the level of abstraction as a compromise between accurate neural modeling and efficient engineering design. In particular, the combination circuit shown in Figure 2(c) uses digital logic to combine spikes; this implementation is not physiologically plausible. In addition, the correlator circuit shown in Figure 3(b) requires non-physiological spike widths for operation, and is another candidate for more realistic models. The inner-hair-cell and spiral-ganglion-cell circuits also lack several important characteristics of their physiological counterparts [4], as does the silicon cochlea [3].

The circuit implements autocorrelation in a straightforward way, using a time-delay element and a spike correlator. A biological implementation may implement autocorrelation differently. For example, a hypothetical correlator neuron could work in the following way. A spike occurring on one input of the correlator neuron would inhibit firing for a set period of time, corresponding to the “delay time” of the correlation. This inhibited time would be followed by a brief period time where a spike on a second input would result in an output spike by the correlator neuron. After this brief time, a single spike on the second input would no longer be sufficient to induce an output spike. Using this “disinhibition” scheme, the neural circuit could be reduced to a single neuron (C. Mead, personal communication).

7. Acknowledgements

I am grateful for C. Mead for providing financial support, laboratory facilities, and inspiration. I also acknowledge the auditory research community associated with Caltech, specifically D. Lyon, M. Konishi, L. Watts, X. Arreguit, E. Corey, and M. Slaney. This work was funded by the Office of Naval Research, the Defense Advanced Research Projects Agency, and the State of California. MOSIS provided chip fabrication.

8. References

- [1] R. P. Lippmann, “Review of neural networks for speech recognition,” *Neural Computation*, vol 1, pp. 1–38, Spring 1989.

- [2] C. Mead, *Analog VLSI and Neural Systems*. Reading, MA: Addison-Wesley, 1989, pp. 3-9.
- [3] R. F. Lyon and C. Mead, "An analog electronic cochlea," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 1119–1134, July 1988.
- [4] J. P. Lazzaro and C. Mead, "Circuit models of sensory transduction in the cochlea," in *Analog VLSI Implementations of Neural Networks*, C. Mead and M. Ismail, Eds. Norwell, MA: Kluwer Academic Publishers, 1989, pp. 85-101.
- [5] J. P. Lazzaro and C. Mead, "Silicon models of auditory localization," *Neural Computation*, vol 1, pp. 47–57, Spring 1989.
- [6] J. P. Lazzaro and C. Mead, "Silicon models of pitch perception," *Proc. Natl. Acad. Sci. USA*, vol 86, pp. 9597–9601, December 1989.
- [7] S. Greenberg, "The ear as a speech analyzer," *J. Phonetics*, vol 16, pp. 139-149, 1988.
- [8] D. O. Kim, "Functional roles of the inner- and outer-haircell subsystems in the cochlea and brainstem," in *Hearing Science*, C. I. Berlin, Eds. San Diego, CA: College-Hill Press, 1984, pp. 241–262.
- [9] P. Dallos, "Response characteristics of mammalian cochlear hair cells," *J. Neurosci*, vol 5, pp. 1591-1608, June 1985.
- [10] E. F. Evans, "Functional anatomy of the auditory system," in *The Senses*, H. B. Barlow and J. D. Mollon, Eds. Cambridge, England: Cambridge University Press, 1982, p. 251.
- [11] S. Shamma, "The acoustic features of speech sounds in a model of auditory processing: vowels and voiceless fricatives," *J. Phonetics*, vol 16, pp. 77-91, March 1988.
- [12] L. Deng, C. D. Geisler, S. Greenberg, "A composite model of the auditory periphery for the processing of speech," *J. Phonetics*, vol 16, pp. 93-108, 1988.
- [13] O. Ghitza, "Temporal non-place information in the auditory-nerve firing patterns as a front end for speech recognition in a noisy environment," *J. Phonetics*, vol 16, pp. 109-123, 1988.
- [14] S. Seneff, "A joint synchrony/mean-rate model of auditory speech processing," *J. Phonetics*, vol 16, pp. 55-76, 1988.
- [15] M. B. Sachs and E. D. Young, "Effects of nonlinearities on speech encoding in the auditory nerve," *J. Acoust. Soc. Am*, vol 68, no. 3, pp. 858-875, September 1980.
- [16] B. Delgutte, "Speech coding in the auditory nerve: II. Processing schemes for vowel-like sounds," *J. Acoust. Soc. Am*, vol. 75, no. 3, pp. 879-886, 1984.
- [17] R. F. Lyon, "Computational models of neural auditory processing," in *Proceedings, 1984 IEEE ICASSP*, San Diego, CA (Mar. 19–21, 1984).

[18] N. Suga, "Auditory Neuroethology and Speech Processing: Complex-Sound Processing by Combination-Sensitive Neurons," in *Auditory Function*, G. M. Edelman, W. E. Gall, W. M. Cowan, Eds. New York: Wiley, 1988, pp. 679–720.

[19] C. Mead, *Analog VLSI and Neural Systems*. Reading, MA: Addison-Wesley, 1989, pp. 173-177.

[20] C. Mead, *Analog VLSI and Neural Systems*. Reading, MA: Addison-Wesley, 1989, pp. 193-203.