Analog Integrated Circuits and Signal Processing, ??, 1–60 (19??)

A Low-Power Wide-Dynamic-Range Analog VLSI Cochlea

RAHUL SARPESHKAR
rahul@physics.bell-labs.com
Department of Biological Computation, Bell Laboratories, Murray Hill, NJ 07974

RICHARD F. LYON
dicklyon@acm.org
Foveonics Inc., Cupertino, CA 95014

CARVER MEAD
donna@pcmp.caltech.edu
Physics of Computation Laboratory, California Institute of Technology, Pasadena, CA, 91125

Received April ??, 1996. Revised April ??, 1996.

Abstract. Low-power wide-dynamic-range systems are extremely hard to build. The biological cochlea is one of the most awesome examples of such a system: It can sense sounds over 12 orders of magnitude in intensity, with an estimated power dissipation of only a few tens of microwatts. In this paper, we describe an analog electronic cochlea that processes sounds over 6 orders of magnitude in intensity, and that dissipates 0.5mW. This 117-stage, 100Hz-to-10Khz cochlea has the widest dynamic range of any artificial cochlea built to date. The wide dynamic range is attained through the use of a wide-linear-range transconductance amplifier, of a low-noise filter topology, of dynamic gain control (AGC) at each cochlear stage, and of an architecture that we refer to as *overlapping cochlear cascades*. The operation of the cochlea is made robust through the use of automatic offset-compensation circuitry. A BiCMOS circuit approach helps us to attain nearly scale-invariant behavior and good matching at all frequencies. The synthesis and analysis of our artificial cochlea yields insight into why the human cochlea uses an active traveling-wave mechanism to sense sounds, instead of using bandpass filters. The low power, wide dynamic range, and biological realism make our cochlea well suited as a front end for cochlear implants.

1. Introduction

The dynamic range of operation of a system is measured by the ratio of the intensities of the largest and smallest inputs to the system. Typically, the dynamic range is quoted in the logarithmic units of decibel (dB), with 10dB corresponding to 1 order of magnitude. The largest input that a system can handle is limited by nonlinearities that cause appreciable distortion or failure at the output(s). The smallest input that a system can handle is limited by the system's input-referred noise floor.

At the *same* given bandwidth of operation, a low-current system typically has a higher noise floor than does a high-current system: The low-current system averages over fewer electrons per unit time than does the high-current system, and, consequently, has higher levels of shot or thermal noise [1]. Thus, it is harder to attain a wide dynamic range in low-current systems than in high-current systems. A low-voltage system does not have as wide a dynamic range as a high-voltage system because of a reduction in the maximum voltage of operation.[1]

Low-power systems have low-current or low-voltage levels; consequently, it is harder to attain a wide dynamic range in low-power systems than in high-power systems. The biological cochlea is impressive in its design because it attains an extremely wide dynamic range of 120dB (at 3kHz), although its power dissipation is only about 14μW. The power dissipation in the biological cochlea has been estimated from impedance calculations to be about $0.4\mu\text{W/mm}\times 35\text{mm} = 14\mu\text{W}$ [2].

The dynamic range of the cochlea at various input frequencies has been measured by psychophysical and physiological experiments [3]. The biological cochlea has a wide dynamic range because it has an adaptive traveling-wave amplifier architecture, and also because it uses a low-noise electromechanical technology.

The electronic cochlea models the traveling-wave amplifier architecture of the biological cochlea as a cascade of second-order filters with corner frequencies that decrease exponentially from 20kHz to 20Hz (the audio frequency range) [4]. The exponential taper is important in creating a cochlea that is roughly scale invariant at any time scale; it is easily implemented in subthreshold CMOS, or in bipolar technology.

Prior cochlear designs have paid little or no attention to dynamic range. The reports do not give their dynamic ranges [4]–[8]. However, we know that low-power cochlear designs that pay no attention to noise or gain control, like our own initial designs, have a dynamic range of about 30dB to 40 dB (1mV to 70mV rms) at the small-signal peak frequency (BF) of a typical cochlear stage. The lower limit of the dynamic range is determined by the input signal level that results in an output signal-to-noise ratio (SNR) of 1. The upper limit of the dynamic range is determined by the input- signal level that causes a total harmonic distortion (THD) of about 4%. Typically, the upper limit is a strong function of the linear range of the transconductance amplifiers used in the cochlear filter.

A single follower-integrator filter in one of our recent designs [9] had a dynamic range of 65dB (0.55mV–1000mV rms) because of the use of a wide-linear-range transconductance amplifier (WLR) [10]. However, even if the first filter in a cochlea has a wide dynamic range, the dynamic range at the output of a typical cochlear stage is reduced by the accumulation and amplification of noise and distortion from stages preceding it. Nevertheless, the constant reduction in the bandwidth of the cochlear stages along the cascade ensures that the total noise or distortion eventually becomes invariant with the location of the cochlear stage: Noise or distortion accumulates along the cascade, but it is also reduced constantly by filtering. However, the asymptotic noise is high enough that, in our design [9], the dynamic range for a cochlear stage with a BF input was only about 46 dB (5mV to 1000mV rms). In that design, the use of nonlinear gain control helped to decrease the small-signal Q with increasing input amplitude, and thus mitigated the effects of distortion; however, the design's filter topology was not low-noise, and the nature of the nonlinear gain-control circuit was such that the circuit increased the noise further. Thus, the effects of noise accumulation and amplification limited our ability to attain a wide dynamic range.

In this paper we describe a cochlea that attains a dynamic range of 61dB at the BF of a typical cochlear stage by using four techniques:

1. The previously described WLR
2. A low-noise second-order filter topology
3. Dynamic gain control (AGC)
4. The architecture of *overlapping cochlear cascades*

In addition, we use three techniques that ensure the presence of a robust infrastructure in the cochlea:

1. Automatic offset-compensation circuitry in each cochlear filter prevents offset accumulation along the cochlea.
2. Cascode circuitry in the WLRs increase the latter's DC gain, and prevent low-frequency signal attenuation in the cochlea.
3. Translinear bipolar biasing circuits provide Qs that are approximately invariant with corner frequency, and allow better matching. Bipolar biasing circuits were first used in cochlear designs by [8].

We shall discuss all of these preceding techniques in this paper.

The organization of this paper is as follows: In Section 2 we discuss the architecture and properties of a single cochlear stage. In Section 3 we discuss the architecture and properties of the cochlea. In Section 4 we compare analog and digital cochlear implementations with respect to power and area consumption. In Section 5, we discuss the relationship between our electronic cochlea and the biological cochlea. In Section 6, we discuss possible applications of the electronic cochlea for cochlear implants. In Section 7, we summarize our contributions.

2. The Single Cochlear Stage

Figure 1 shows a schematic for a singe cochlear stage. The arrows indicate the direction of information flow (input to output). The second-order filter (SOS) is composed of two WLR amplifiers, two capacitors, and offset-compensation circuitry (LPF and OCR). The corner frequency $1/\tau$ and quality factor Q of the filter are proportional to $\sqrt{I_1 I_2}$ and $\sqrt{I_1/I_2}$, respectively, where I_1 and I_2 are the bias currents of the WLR amplifiers. The tau-and-Q control circuit controls the values of the currents I_1 and I_2 such that the value of $1/\tau$ depends on only the bias voltage V_T, and the small-signal value of Q depends only on the bias voltage V_Q. An

AGC correction current I_A attenuates the small-signal value of Q at large-signal levels in a graded fashion.

The inner-hair-cell circuit (IHC) rectifies, differentiates, and transduces the input voltage to a current I_{hr}. The voltage V_A controls the value of an internal amplifier bias current in the IHC. The voltage V_{HR} controls the transduction gain of the IHC. The peak detector (PD) extracts the peak value of I_{hr} as a DC current I_{pk}. The current I_{pk} becomes the AGC correction- current input (I_A) to the tau-and-Q control circuit. The bias voltage V_{PT} determines the time constant of the peak detector, and thus the response time of the AGC. The peak detector is designed such that it can respond to increases in input intensity within one cycle of a sinusoidal input at V_{in} ; its response to decreases in input intensity is much slower and is determined by V_{PT}.

The offset-compensation circuit is composed of a lowpass filter (LPF) whose time constant is determined by V_{OT}. The LPF extracts the DC voltage of the filter's intermediate node, and compares this voltage with a global reference voltage V_{RF} in the offset-correction block (OCR). The OCR applies a correction current to the intermediate node to restore that node's voltage to a value near V_{RF}. The DC voltage of the output node is then also near V_{RF}, because the systematic offset voltage of a WLR amplifier is a small negative voltage. The maximal correction current of the OCR scales with the bias current I_1; the bias voltage V_{OF} controls the scaling ratio. Since the restoration is performed at every cochlear stage, the output voltage of each stage is near V_{RF}, and offset does not accumulate across the cochlea. If there were no offset adaptation, a systematic offset voltage in any one stage would accumulate across the whole cochlea.

Since the gain-control topology is feedforward, rather than feedback, we avoid instabilities or oscillations in the Q. However, since raising Qs lowers the DC voltage, and the DC voltage does have a mild influence on the Q, the DC and AC output-voltage dynamics are weakly dependent on each other.

We shall now describe the details of each of the circuits in Figure 1. In Section 2.1 we discuss the WLR circuit. In Section 2.2, we describe the offset-adaptation circuit. In Section 2.3, we examine the filter topology. In Section 2.4 we present the translinear tau-and-Q control circuit. In Section 2.5, we describe the circuits in the IHC and PD blocks. In Section 2.6, we discuss the overall properties of an entire cochlear stage.

2.1. The WLR

The WLR has been described in great detail [10]. The version of the WLR that we use in our cochlea, however, has been slightly modified, so we shall describe it briefly. Figure 2 shows the circuit of the transconductance amplifier. The inputs v_+ and v_- are the wells of the W transistors; we use the well, instead of the gate, to lower amplifier transconductance and consequently to widen the linear range of the amplifier. The linear range is further widened through the novel technique of gate degeneration via the GM transistors, and through the technique of bump linearization via the B transistors [10]. The GM–M mirrors are attenuating, with a 1:3 ratio, to avoid parasitic-capacitance effects in the differential pair. The CP and CN transistors function as cascode transistors; they ensure that the DC gain of the amplifier is high such that there is no significant low-frequency attenuation in the cochlea (See Section 3.1 for further details.). We use CP and CN transistors on both arms of the amplifier to avoid any systematic offsets. The CP and CN transistors do not alter the noise performance of the amplifier, since their noise currents contribute little to the output current of the amplifier. The pFET M transistors implement current mirrors. The bias current of the amplifier I_B is provided by bipolar transistors in the tau-and-Q biasing circuit. The extra mirror necessary to convert NPN bipolar collector currents to pFET bias currents does not alter the noise performance of the amplifier. The offset-correction circuit adds two correction currents at the V_{ol} and V_{or} nodes. In the filter of Figure 1, only the left amplifier has correction-current inputs from the OCR.

2.2. The Offset-Adaptation Circuit

Figure 3 shows the offset-adaptation circuit. The LPF is a simple 5-transistor nFET OTA-C filter operated in the subthreshold regime. The DC value of the V_1 input is extracted by the LPF and is compared with V_{RF} in the pFET differential pair. The currents in the arms of the differential pair are steered via mirrors to the V_{ol} and V_{or} inputs of the left WLR in Figure 1, such that the sign of the feedback is negative. The cascoding of the offset-correction currents prevents the offset-correction circuitry from degrading the DC gain of the amplifier. The V_B gate voltage of the left WLR, as revealed in Figure 2, is also the V_B gate voltage in the offset-adaptation circuit, such that the offset-correction current scales with the bias current of the WLR. The scaling ratio is controlled by the source control V_{OF}.

The value of V_{OF}, which determines the loop gain of the offset-adaptation loop, and the value of V_{OT} determine the closed-loop time constant of the offset-adaptation loop. A high loop gain speeds up the closed–loop response such that the influence of low-frequency inputs is adapted away. Thus, there is a tradeoff between strong offset control (high loop gain), which implies some ringing and overshoot as well, and coupling of low-frequency inputs into the cochlea. Since our lowest-frequency input is 100Hz, we have been able to maintain a good control of the DC offset in the cochlea without attenuating any frequencies of interest.

Note that the offset-correction circuitry can correct offsets to only a resolution that is limited by its own offsets. Since we use subthreshold circuitry with small linear ranges in the OCR and LPF, these offsets are on the order of 5mV to 15mV; they are small compared with the 1V linear range of the WLR. The offset-correction scheme would have been less effective if we had used other WLRs to sense and correct the offset of the filter WLRs. In the latter case, since the offset of a WLR scales with its linear range, the resolution of the offset-correction circuitry would typically have been a significant fraction of 1V.

Note that our offset-compensation circuitry does not require any floating gates and the assosciated need for high-voltages, and high-voltage circuits. Further, unlike in some floating-gate circuits, the time constant of the offset-adaptation circuit may be tuned to be in the 10ms.–1sec. range; such values ensure that the offset adaptation across an entire cochlea is not excessively slow.

2.3. The Second-Order Filter

Figure 4 shows representations of our second-order filter. The block-diagram form of part (b) is convienient for doing noise calculations. For the purposes of doing noise calculations we list 12 algebraic relationships:

$$g_{m1} = \frac{I_1}{V_L}. \tag{1}$$

$$g_{m2} = \frac{I_2}{V_L}. \tag{2}$$

$$\tau_1 = \frac{C_1}{g_{m1}}. \tag{3}$$

$$\tau_2 = \frac{C_2}{g_{m2}}. \tag{4}$$

$$\tau = \sqrt{\tau_1\tau_2}. \tag{5}$$

$$I_1 = \frac{C_1 V_L}{\tau_1}. \tag{6}$$

$$I_2 = \frac{C_2 V_L}{\tau_2}. \tag{7}$$

$$Q = \sqrt{\frac{\tau_2}{\tau_1}}. \tag{8}$$

$$\tau_1 = \frac{\tau}{Q}. \tag{9}$$

$$\tau_2 = \tau Q. \tag{10}$$

$$C_R = \frac{C_2}{C_1}. \tag{11}$$

$$C = \sqrt{C_1 C_2}. \tag{12}$$

The currents I_1 and I_2 are the bias currents of the amplifiers, V_L is the linear range of these amplifiers, and the transconductance g_{mi} of amplifier i is given by $g_{mi} = I_i/V_L$. The noise source v_{ni}, shown in Figure 4(b), represents the input-referred noise per unit bandwidth of amplifier i. From [10], we know that

$$v_{ni}^2 = NqV_L^2/I_i, \tag{13}$$

where N is the effective number of shot-noise sources in the amplifier, and q is the charge on the electron. For our amplifiers, N is typically 4.8, whereas the amplifiers reported in [10] have N = 5.3.

From Figure 4(b) and the preceding algebraic relationships, it may be shown that the output noise per unit bandwidth v_{no} is given by

$$v_{no} = \frac{v_{n1} + v_{n2}(\tau s/Q)}{\tau^2 s^2 + \tau s/Q + 1}, \tag{14}$$

$$v_{no}^2 = \frac{v_{n1}^2 + \left|\tau^2 s^2/Q^2\right| v_{n2}^2}{\left|\tau^2 s^2 + \tau s/Q + 1\right|^2}. \tag{15}$$

In Eq. (15) we have used the fact that the noise sources v_{n1} and v_{n2} are uncorrelated, so there are no cross terms of the form $v_{n1}v_{n2}$. From the algebraic relationships of Eqs. (1) through (12), by substituting $s = \omega\tau$ and $\omega = 2\pi f$, and by using the normalized frequency

$x = \omega\tau$, we can show that

$$v_{no}^2(f)df = \frac{(NqV_L/QC_1)\left(1+\frac{\omega^2\tau^2}{C_R}\right)}{(1-\omega^2\tau^2)^2+\frac{\omega^2\tau^2}{Q^2}}df, \quad (16)$$

$$v_{no}^2(x)dx = \left(\frac{NqV_L}{2\pi QC_1}\right)|H(x)|^2(1+\frac{x^2}{C_R})dx, \quad (17)$$

where $H(x)$ represents the normalized transfer function of a second-order filter,

$$H(x) = \frac{1}{1-x^2+jx/Q}, \quad (18)$$

and j is the square root of -1. To get the total noise over all frequencies, we integrate the LHS and RHS of Eq. (17) from 0 to ∞. It can be shown by contour integration that

$$\int_0^\infty |H(x)|^2 dx = \frac{\pi}{2}Q, \quad (19)$$

$$\int_0^\infty x^2|H(x)|^2 dx = \frac{\pi}{2}Q. \quad (20)$$

$$(21)$$

The total output noise over all frequencies $< v_{no}^2 >$ is then given by

$$< v_{no}^2 > = \frac{1}{2\pi}(\frac{NqV_L}{QC_1})(\frac{\pi}{2}Q+\frac{\pi}{2}Q(\frac{1}{C_R})), \quad (22)$$

$$= \left(\frac{NqV_L}{4C_1}\right)(1+\frac{1}{C_R}), \quad (23)$$

$$= \left(\frac{NqV_L}{4C}\right)(\sqrt{C_R}+\frac{1}{\sqrt{C_R}}). \quad (24)$$

Note that the total output noise is independent of Q for this topology: The noise per unit bandwidth scales like $1/Q$, and the integration of the noise over all bandwidths scales like Q, so that the total output noise is independent of Q. These relationships are reminiscent of an LCR circuit where the total output current noise depends on only the inductance, the total output voltage noise depends on only the capacitance, and neither of these noises depends on R; only the noise per unit bandwidth and the Q of the LCR circuit are influenced by R. In fact, it can be shown that this topology has a transfer function and noise properties that are similar to those of an LCR circuit if we make the identifications

$$\tau_1 = L/R, \quad (25)$$

$$\tau_2 = RC. \quad (26)$$

For a given value of C (the geometric mean of C_1 and C_2), the total output noise is minimized if C_R = 1—that is, if $C_1 = C_2$.

Figure 5 shows the noise spectral density versus frequency and the total integrated noise at the output over all frequencies. The data were obtained from a test chip that contained a second-order filter with amplifiers identical to those in our previous report [10]. The lines are fits to theory. As we expect, the total integrated noise is quite constant with Q. The parameters used in the fits were N = 5.3, V_L = 1V, C_R = 1.43, $q = 1.6 \times 10^{-19}$, and C = 697fF. The values of N and V_L were obtained from measurements on our transconductance amplifiers [10]. The value of C_R was obtained from least-squares fits to the data. We obtained the value of C by having the noise measurements be consistent with values of C expected from the layout.

In filter topologies that have been used in prior cochlear designs e.g. [4] or [9], the noise per unit bandwidth increases with Q: The Q is obtained through the addition of positive-feedback currents. These currents contribute additional shot noise and thus increase the noise per unit bandwidth; in these topologies, the integrated noise over all frequencies also increases with Q, so both factors contribute to the increase in noise with Q. Although we have performed extensive experimental and theoretical analysis of noise in these filter topologies as well, we shall present only key findings here: For the topology presented in [4], at Qs near 2.5, the rms noise power at the output is 9 times higher than it is for our topology. For Qs near 0.707, the rms noise power at the output is about 0.8 times lower than it is for our topology. For Qs near 1.5, which is where we typically operate, the rms noise power at the output is about 2 times higher than it is for our topology. The effects of increased noise per unit bandwidth in a single second-order filter are greatly amplified in a cochlear cascade. Factors of 2 in noise reduction in a single stage can make a significant reduction in the output noise of a cochlear cascade. Thus, using our filter topology contributes significantly to reducing noise in a cochlear cascade.

Although the noise properties of the filter of Figure 4 are superior to those of other second-order topologies, this filter's distortion at large amplitudes is significantly greater, especially for Qs greater than 1.0: Distortion arises when there are large differential voltages across the transconductance-amplifier inputs in a filter. The feedback to the first amplifier of Figure 4 arises from V_2, rather than from V_1, in contrast to the topology of [4]. Consequently, the accumulation of phase shift from two amplifiers, as opposed to that from one amplifier used in earlier topologies, causes greater differential voltages and greater distortion in the first amplifier. Also, the transfer function of the intermediate node V_1 is such that the magnitude of voltage at this node is greater than that in other topologies for Qs greater than 1.0. Consequently, the differential voltage across the second amplifier is larger, and the distortion from the second amplifier is also greater.

It is instructive to find the largest input signal at the BF of a filter for which the total harmonic distortion (THD) is about 3%–5%. The amplitude of this signal, v_{mx}, is a good measure of the upper limit of dynamic range for a filter, in the same way that the input-referred noise is a good measure of the lower limit of dynamic range. Figure 6 shows the rms amplitude v_{mx} at a BF of 140Hz for the filter of Figure 4. We observe that, as the Q increases, the distortion increases, and the value of v_{mx} falls. The data were obtained for a THD level of 3.3% (30dB attenuation in intensity). The data were *empirically* fit by the equation

$$v_{mx}(Q) = 128 - 161 \ln\left(\frac{Q}{1.75}\right). \qquad (27)$$

The preceding discussion illustrates why an AGC is essential for attaining a wide dynamic range with our filter topology: The noise properties of the topology are favorable for sensing signals at small amplitudes, and with high Qs. However, when the signal levels are large, if the distortion is to be kept under control, the Qs must be attenuated. The AGC ensures that the Qs are large when the signal is small, and are small when the signal is large.

2.4. The Tau-and-Q Control Circuit

In Figure 7, we make the following definitions:

$$I_\tau = I_s e^{V_T/2U_T}, \qquad (28)$$
$$Q_0 = e^{V_Q/2U_T}, \qquad (29)$$

where $U_T = kT/q$ is the thermal voltage, and I_s is the bipolar preexponential constant. The current I_A is a place holder for an AGC correction current from the IHC and peak-detector circuit, and I_1 and I_2 are output currents that bias the first and second amplifiers of Figure 4, respectively. A simple translinear analysis and the solution of the quadratic equation reveal that, if we define η to be a normalized AGC correction current, according to

$$\eta = \left(\frac{I_A}{2I_\tau/Q_0}\right), \qquad (30)$$

then

$$Q = \sqrt{\frac{I_1}{I_2}}, \qquad (31)$$
$$= Q_0\sqrt{\left(\frac{\sqrt{1+\eta^2}-\eta}{\sqrt{1+\eta^2}+\eta}\right)}. \qquad (32)$$

Independent of the value of I_A, the translinear circuit always ensures that

$$\sqrt{I_1 I_2} = I_\tau. \qquad (33)$$

Thus, it is an effective tau-and-Q biasing circuit for the filter in Figure 4, since it ensures that the AGC affects the Q but not the corner frequency of the filter. If we let

$$\theta = \arctan \eta, \qquad (34)$$

then trigonometric manipulations of Eq. (32) reveal that

$$Q = Q_0 \tan\left(\frac{\pi}{4} - \frac{\theta}{2}\right). \qquad (35)$$

If there is no AGC correction current, then $\theta = 0$ and $Q = Q_0$. In the limit of an infinite AGC correction current, $\theta/2 = \pi/4$ and $Q = 0$.

Figure 8(a) shows the corner frequency of the filter in Figure 4 as a function of the bias voltage V_T. As we expect from Eq. (28) and (33), and from the equations of the filter (Eqs. 1 to (12)), the corner frequency is an exponential function of the bias voltage V_T. The exponential preconstant yields a thermal voltage of 26.7mV, which is fairly close to the expected thermal voltage of 26mV at a room temperature of 300K.

Figure 8(b) shows the Q of the filter in the absence of any AGC correction current. As we expect from Eq. (29) and Eq. (35) with $\eta = 0$ (no AGC current), the

Q is an exponential function of the bias voltage V_Q. The exponential preconstant yields a thermal voltage of 26.3mV, which is fairly close to the expected thermal voltage of 26mV at a room temperature of 300K.

2.5. The Inner Hair Cell and Peak-Detector Circuits

Figure 9 shows the IHC and PD circuits. The amplifier in the IHC is a simple 5-transistor nFET OTA with a fairly high gain (500 to 1000). The bias current of the OTA is determined by the voltage V_A. The bias current should be sufficiently high that the dynamics of the node V_h are much faster than the dynamics of the node V_n, for all input frequencies and amplitudes of interest. Since the OTA is connected in a follower configuration, the voltage V_n is very nearly a copy of V_{in}, except for very weak signals, where the bipolar transistor BA or the MOS transistor PA are not sufficiently turned on. In practice, the signals or noise at the cochlear output taps are sufficiently high that BA or PA may be assumed always to be sufficiently turned on. When V_{in} or V_n are rising, the capacitor C_{HR} is charged primarily by the bipolar transistor BA. When V_{in} or V_n are falling, the capacitor C_{HR} is discharged primarily by the MOS transistor PA. Thus, during the phases of the signal when the derivative of the signal is negative, the current I_{hr} is an amplified copy of $C_{HR}dV_{in}/dt$. The amplification factor is given by $\exp(V_{HR}/U_T)$. Thus, the IHC differentiates, rectifies, amplifies, and transforms the input voltage V_{in} into an output current I_{hr}. The use of a bipolar transistor BA and a source-connected MOS transistor PA ensure that back-gate effects do not reduce the circuit's driving capability for large-signal inputs.

The output current I_{hr} is fed into the peak detector. The peak detector consists of a slow source follower, composed of PF, PT, and $C_{\rm PT}$, and the feedback transistor PI. The transistor PO outputs a copy of the current in PI as I_{pk}. The source follower can follow descending signals in V_f rapidly because of the exponential dependence of the current of PF on its gate voltage. However, the voltage V_{PT} is set near V_{DD} so that the current source formed by the transistor PT is slow in charging the capacitor C_{PT}; consequently, during ascending signals in V_f, the voltage V_s is slow to respond. Because of the fedback nature of the circuit, and the asymmetry in the time constants of the source follower, V_s will equilibrate at a value such that the average current through PI is slightly below the peak value of I_{hr}. As I_{hr} alternately reaches its peak and moves below that peak, the voltage V_f will undergo large swings due to the high gain of the input node of the peak detector. In contrast, the voltage V_s will have only small variations from its DC value; they constitute the ripple of the peak detector.

Figure 10 shows the waveforms V_{in}, V_n, V_h, V_f, and V_s. The labeled voltages in the figure indicate the DC voltage values that correspond to the horizontal location of the arrow. As we expect, V_{in} and V_n are very nearly equal to each other. The voltage V_h undergoes abrupt transitions during changes in the sign of the input derivative; these changes correspond to a transition from BA being turned off to PA being turned on or vice versa. The voltages V_f and V_s in the peak detector undergo rapid downward transitions that are phase locked to the downward-going zero crossings of the input waveform where the peak value of I_{hr} occurs. The upward transitions in V_f and V_s are slow because of the sluggishness of the current-source transistor PT. The data were taken with V_{in} being a 102mV rms input at 1kHz, with V_A = 1.0V, with V_{PT} = 4.039V, V_{DD} = 5.0V, and with V_{HR} = 100mV. Typically, we operate V_{PT} near 4.25V, which results in no discernible ripple in V_s, but these data were taken specifically to illustrate better the workings of the peak detector. The transistor PT was fabricated as a poly2 transistor. Thus, at the same current level, the bias voltages on V_{PT} are higher than those corresponding to bias voltages on a poly1 transistor.

From the preceding discussion, we expect that the value of I_{pk} will be near the peak value of $C_{HR}dV_{in}/dt$ amplified by the factor of $\exp(V_{HR}/U_T)$. Thus, if the input amplitude were given by

$$V_{in} = a_{in} \sin{(2\pi f_{in} t)}, \tag{36}$$

then the value of I_{pk} would be given by

$$I_{pk} = 2\pi f_{in} C_{HR} a_{in} e^{V_{HR}/U_T}. \tag{37}$$

In conformance with Eq. (37), Figure 11 shows that the response of I_{pk} is linear with the amplitude and with the frequency of the input. The data were taken for V_A = 1.0V, and V_{PT} = 4.3V. The experimental slopes for Figure 11(a) and Figure 11(b) yielded values for C_{HR} = 335fF and C_{HR} = 313fF, respectively. However, the linear fits to the data reveal that an offset in amplitude of about 77.5mV rms in the case of Figure 11(a), and an offset in frequency of about 276Hz in the case of Figure 11(b), needs to be subtracted from a_{in} or f_{in}, respectively. These offsets imply that there is a minimum amount of input current I_{hr} that is required for the peak detector to output a current I_{pk}. Through experimentation, we have found that this minimum value scales approximately linearly with frequency such that the offset for a_{in} always lies somewhere in the 50 to 100mV rms region (for a V_{HR} of about 100mV). At this time, we do not have a good explanation of what causes these offsets; we suspect that they are due to the short-channel length and small Early Voltage of transistor PI.

Figure 12 shows that the relationship between I_{pk} and V_{HR} is described by an exponential, as Eq. (37) predicts. The thermal voltage U_T was determined to be around 29.9mV. This voltage is somewhat higher than the 26mV expected from theory. The data were taken with V_{PT} = 4.30V, and V_A = 1.15V.

The current I_{pk} is mirrored by the bipolar transistors BP and BO in Figure 9 to function as the AGC correction current I_A in Figure 7. From Eqs. 1 to (12), we know that I_τ is given by $2\pi f_c C V_L$, where $f_c = 1/\tau$ is the corner frequency (CF) of the filter. Thus, η in Eq. (30) is given by

$$\eta = \frac{I_A}{2I_\tau/Q_0}, \tag{38}$$

$$= Q_0 e^{V_{HR}/U_T}\left(\frac{f_{in}}{f_c}\right)\left(\frac{C_{HR}}{2C}\right)\left(\frac{a_{in}}{V_L}\right). \tag{39}$$

Thus, the voltage V_{HR} serves to strengthen or weaken the normalized AGC correction factor η.

2.6. The Properties of an Entire Cochlear Stage

Figure 13 shows the frequency-response characteristics of the filter of Figure 4 for different input amplitudes. In the absence of an AGC, large input amplitudes generate large amounts of distortion in the filter; thus, in Figure 13(a), it was impossible to obtain smooth frequency responses beyond an input rms amplitude of 250mV. In contrast, in Figure 13(b), we could easily obtain smooth frequency responses up to (and even beyond) input amplitudes of 1V rms, because of the presence of an AGC. If the frequency-response curves are fitted with the transfer function of a second-order section,

$$H(s) = \frac{1}{\tau^2 s^2 + \tau s/Q + 1}, \tag{40}$$

then we find that the CF ($1/\tau$) is reduced with input amplitude, and the Q is reduced by the AGC as well. In Figure 13(b), the CF is reduced from 147Hz at small signals to 112Hz at large signals; the Q is reduced from 2.22 at small signals to about 0.52 at large signals. These numbers are valid for V_{HR} = 65mV, and V_{PT} = 4.25V. Given that we designed Eq. (33) in our translinear biasing circuit to keep the CF constant, it may seem surprising that the CF changed with input amplitude. However, we must realize that we are fitting the frequency-response curves of a nonlinear system with a linear approximation given by Eq. (40) at each rms input amplitude. According to Eqs.(39) and (32), for the same input rms amplitude, the "Q" is lower at high frequencies than at low frequencies. The frequency dependence of the Q results in disproportionately more attenuation of the input at high frequencies than at low frequencies, such that the CF, as measured by the amplitude curves, appears to shift.

If we plot the CF_{90}—that is to say the frequency at which the the second-order filter has a phase lag of 90 degrees—versus rms input amplitude, then the data of Figure 14 reveal that CF_{90} is approximately constant. At low input amplitudes, the AGC has no effect, because I_{pk} provides no correction until the input amplitude is above a certain threshold, as we discussed in Section 2.5. Even if there were no offset, the AGC correction in this regime would be small. Thus, the system is linear at low amplitudes. Consequently, at these amplitudes, the CF_{90} is identical with $1/\tau$ and with the CF measured by gain curves. Since the AGC is designed not to affect the parameter τ, the CF_{90} remains approximately invariant with input amplitude, even at high amplitudes where the AGC is active. In fact, Figure 14 shows that a strong AGC (higher values of V_{HR}) improves the constancy of the CF_{90} with amplitude, because it prevents static nonlinear shifts in CF_{90} that increase at high Qs. The CF_{90} is the frequency near which center-surround schemes, e.g., those that perform spectral extraction on cochlear outputs for use in implants [11], generate their maximum output. Thus, the fact that the CF_{90} is approximately invariant with amplitude makes our AGC cochlea attractive as a front end for center-surround postprocessing.

Figure 15(a) shows data for Q versus a_{in} measured for three different values of V_{HR} for the second-order filter. From Eqs. (30), (34), and (35) we would expect the curves to be fit by a function of the form

$$Q = Q_0 \tan\left(\frac{\pi}{4} - \frac{\arctan\left(Ga_{in}\right)}{2}\right). \quad (41)$$

However, from the discussion of Section 2.5, we know that a_{in} in Eq. (41) should be replaced by 0 below some threshold value a_0, and by $a_{in} - a_0$ above this threshold value. The fits in Figure 15(a) are functional fits to Eq. (41) with the free parameter G, and the additional free parameter a_0. For V_{HR} = 3mV, 32mV, and 65mV, we found $G = 1.93, 2.5, 4$, and $a_0 = 0.164, 0.095, 0.06$, respectively; Q_0 was 2.05 for all curves, a_{in} and a_0 are in units of V. We took data by measuring the gain of the filter at f_{in} = $f_c = CF_{90} = 1/\tau$, which, for a second-order filter, is Q. Figure 15(b) plots the output amplitude, $Q(a_{in}) \times a_{in}$, rather than $Q(a_{in})$, at this frequency. We observe that, before the AGC turns on ($a_{in} < a_0$), the relationship between the input and output amplitudes is linear. After the AGC turns on ($a_{in} > a_0$), the relationship between the output and input amplitudes is compressive, although not as compressive as theory would predict. Since a_0 is large for small values of V_{HR}, the range of linearity is large for small values of V_{HR}.

Figure 15 suggests that, at large amplitudes, the static nonlinearities in the filter increase the Q slightly. Since the Q of the filter is given by $\sqrt{\frac{\tau_2}{\tau_1}}$, we deduce that the static nonlinearity is causing τ_2 to increase faster with a_{in} than τ_1; this deduction is in accordance with the intuition that the second amplifier in Figure 4 is subject to greater differential voltages, and, consequently, to more saturation and slowing than the first amplifier. One way to avoid, or even to reverse, the nonlinear shift toward higher Qs is to have the linear range of the first amplifier be smaller than the linear range of the second amplifier.

The nature of Eq. (41) is such that, independent of the value of G, $Q(a_{in})a_{in}$ is a monotonically increasing function of a_{in}. This property guarantees that the input–output curve at the BF of a cochlear stage is always monotonically increasing, as confirmed by the data of Figure 15.

Figure 16 shows that the harmonic-distortion levels at 1V rms with an AGC are comparable with the harmonic-distortion levels at 250mV rms without an AGC. The AGC data were taken with V_{hr} = 65mV. Figure 17 shows that a strong AGC (large value of V_{HR}) reduces harmonic distortion due to the lowering of Q.

Figure 18 illustrates the dynamics of Q adaptation: The stimulus is a pure tone at the BF of the filter that turns on suddenly after a period of silence, persists for a while, and then abruptly decreases in intensity to a quieter tone. At the onset of the tone, a transient response is seen at the output. The transient causes the peak detector to overadapt instantly within one cycle. The overadaptation is corrected by the slow capacitive charging of the peak-detector current source, which restores the Q, and thus the output amplitude, to an equilibrium value. When the tone transitions from loud to soft, the initial response to the soft tone is moderate due to the low Q caused by adaptation to the preceding loud tone. Eventually, the slow capacitive charging of the peak-detector current source restores the Q, and thus the output amplitude, to an equilibrium value.

3. Properties of the Cochlea

In this section, we shall discuss the properties of the cochlea. We shall begin with a discussion of low-frequency attenuation because the discussion will motivate the introduction of our overlapping-cascades architecture.

3.1. Low-Frequency Attenuation

If the open-loop gains of the amplifiers in Figure 4 are A_1 and A_2, then we can show that we obtain the low-frequency gain of the filter of Figure 4 by simply replacing $\tau_1 s$ and $\tau_2 s$ with $1/A_1$ and $1/A_2$ in the transfer function. Thus, from Eq. (40) the low-frequency gain H_0 is given by

$$H_0 = \frac{1}{1 + \frac{1}{A_1} + \frac{1}{A_1 A_2}} \quad (42)$$

$$\approx \frac{A_1}{A_1 + 1}. \quad (43)$$

Although H_0 is very close to 1, it is not exactly 1. A low-frequency input that travels through M stages of a cochlea will suffer a net attenuation, H_M, given by

$$H_M = \left(\frac{A_1}{A_1 + 1}\right)^M, \quad (44)$$

$$\approx e^{-M/A_1}, \tag{45}$$

where the exponential approximation is valid if A_1 is a large number. Now, $A_1 = 2V_0/V_L$, where V_L is the linear range of the amplifier and V_0 is its effective Early voltage at the output of the amplifier [10]. In the 1.2μm *n*well Orbit MOSIS process used to fabricate our circuits, V_0 is around 20V for wide-linear-range amplifiers that do not have cascode transistors at the output; V_L = 1V, and A_1 is about 40. Thus, we can expect an attenuation of almost one e-fold across a 39-stage cochlea built with cascodeless amplifiers.

Figure 19 shows the low-frequency attenuation of a 40Hz 50mV input to a 39-stage cochlea tuned from 900Hz to 100Hz for different values of the parameter A_1. For these experiments, we operated the cochlea with a very low Q (V_Q = -80mV) so that we could focuss on just the effects of low-frequency attenuation. We varied the value of A_1 by varying the bias of the cascode transistors V_{CN} and V_{CP}, in the amplifier of Figure 2. We explored the effects of turning off the cascode transistors by biasing them as switches. Thus, to turn off the CN cascode transistors, we would set V_{CN} to 5V; to turn off the CP cascode transistors we would set V_{CP} to 0V. Figure 19 shows the low-frequency attenuation for the four cases of both cascodes on, both cascodes off, only N cascodes off, or only P cascodes off. We observe from the data that the P cascodes are more helpful in reducing low-frequency attenuation than are the N cascodes, because the pFETs in our process have a lower Early voltage than do the nFETs. With both cascodes on, a 39-stage cochlea has a net attenuation that is less than 0.8. We normally operate the cochlea with both cascodes on, with V_{CN} = 1.2V, and with V_{CP} = 3.8V. These bias values permit operation of our amplifiers over the entire frequency range of the cochlea without any saturation effects for input rms amplitudes that exceed 1V rms.

The attenuation of the gain of signals at other frequencies is by the same factor H_M. In contrast, the output noise (or distortion) at a cochlear tap is accumulated through addition over successive stages, as shown in Figure 20. The noise that is added at the input is attenuated by the same amount as the signal, but the amounts of noise that are added at stages successively closer to the output tap of interest are attenuated by successively smaller amounts. Thus, the output SNR is degraded by low-frequency attenuation.

To limit the degradation of the SNR of the cochlea through low-frequency attenuation, and noise-and-distortion accumulation, we use the architecture of overlapping cascades shown in Figure 21. Rather than having one large cochlea, we use a few small cochleas whose frequency ranges overlap by one octave. All such cochleas process the input in parallel. The filters in the overlapping octave serve to mimic the effects of the infinite cascade prior to the stages of interest; the outputs of these filters are not used. Since most of the effect of the infinite cascade occurs within an octave of the corner frequency of a cochlear stage, we do not sacrifice much in the way of cochlear modeling, but we do gain significantly in limiting our SNR degradation. In general, the amount of overlap between cochleas, and the number of stages per cochlea can be varied to suit the nature of the cochlear application.

Although the thermal noise in an infinite cascade converges to an equilibrium where noise accumulation is matched by noise filtering, the 1/f noise in an infinite cascade does not converge and continues to grow in the cascade. The 1/f noise is significant for only those high-frequency cochlear stages that have amplifiers with large bias currents [10]. The overlapping-cascades architecture helps to limit the accumulation of 1/f noise.

A cochlear cascade that is composed of all-pole second-order filters overestimates the group delay of the biological cochlea. The overlapping-cascades architecture also helps to reduce the group delay of the silicon cochlea.

The architecture of overlapping cascades may be viewed as a hybrid of an architecture that has many parallel filters in a filter bank and of one that has one filter cascade with all the filters in serial.

The cochlea that we discuss in this paper was built out of three 39-stage overlapping cochlear cascades: The low-frequency cochlear cascade was tuned to operate in the 100Hz to 900Hz region. The mid-frequency cochlear cascade was tuned to operate in the 450Hz to 4050Hz region. The high-frequeny cochlear cascade was tuned to operate in the 2000Hz to 18,000Hz region. Thus, each of the cochlear cascades had about 11.2 filters per octave, ensuring a fairly sharp cochlear rolloff slope. The Qs of the cochleas were tuned to be approximately 1.5. The voltage gradients in V_T corresponding to the three frequency gradients of the low-frequency, mid-frequency, and high-frequency cochlear cascades were 1.040 to 0.9V, 1.130 to 0.990V, and 1.210 to 1.070V respectively. The value of V_Q that was suitable for operating all three cochleas was

-52mV. For the remainder of the paper, we shall focuss on the operation of the low-frequency cochlear cascade, which we shall call the cochlea. The operation of the other cochlear cascades follows by straightforward generalization. The other parameters that we used for operating our cascades were V_{OL} = 4.93V, V_{HR} = 120mV, V_{PT} = 4.25V, V_{CN} = 1.2V, V_{CP} = 3.8V, V_{OT} = 0.3V, V_{RF} = 3.0V, and the DC value of V_{in} = 3V. To conserve power, we operated V_A at 0.995V in the low-frequency cochlea, at 1.05V in the mid-frequency cochlea, and at 1.15V in the high-frequency cochlea. It is possible to reduce the power dissipation even further by having a tilt in the values of V_A in each cochlea. Through experimentation, we found that V_Q = -44mV, -52mV, and -65mV yielded the best performance for the low-frequency, mid-frequency, and high-frequency cochleas, respectively. We could also speed up the gain adaptation in the mid-frequency and high-frequency cochleas by setting V_{PT} in the 4.10V to 4.15V range. We used standard shift-register and clocking circuitry to multiplex the outputs from the different cochlear taps onto a common output tap.

3.2. *Offset Adaptation*

Figure 22 shows the DC output voltage across the cochlea as we scan from tap 1 to tap 39. In the absence of any offset adaptation (V_{OF} = 4.76V), each cochlear stage has a systematic negative offset of about 42mV; by 39 stages the DC output voltage has dropped from 3V to 1V. As we strengthen the offset adaptation by raising the value of V_{OF}, the offset degradation improves. At 4.96V, there is little offset accumulation, and there is an almost flat DC response across the whole cochlea. Typically, we operate the cochlea at V_{OF} = 4.93V and tolerate some offset in return for reduced ringing in the offset-adaptation loop, and for a lower adaptation corner frequency.

3.3. *Frequency Response*

Figure 23(a) shows the frequency response of the cochlea at different input amplitudes ranging from 5mV to 1000mV rms at cochlear tap 30. The adaptation in Q with increasing input amplitude is evident. Figure 23(b) plots the gain versus frequency such that the curve with the highest gain corresponds to the lowest input amplitude of 5mV. The gain adapts from about 30 for the 5mV rms case to about 0.7 at 1000mV rms. Figure 24 shows that the output is approximately linear in the input at frequencies before the BF, is compressive at the BF, and is even more compressive after the BF. These compression characteristics are seen in the biological cochlea as well [12]; they arise because of the accumulated effects of gain adaptation over several cochlear stages.

Figure 25(a) illustrates that the harmonic distortion is greatest about one octave before the BF. This effect occurs because the second-harmonic distortion is amplified by the high gain at the BF when the input frequency is 1 octave before the BF. When the input frequency is at the BF, the second-harmonic distortion drops sharply because 1 octave after the BF there is great attenuation. These effects imply that nonlinearities in the cochlea cause masking in the perception of harmonic frequencies; that is, the threshold for the detection of a 2f tone is higher in the presence of a 1f tone than in the absence of one. Psychophysical experiments reveal this effect in humans as well [13].

Figure 25(b) illustrates the growth and filtering of harmonic distortion as the signal travels through the cochlea. The input is a 1V rms signal with frequency 162Hz that corresponds to the BF at tap 30. As the signal travels from tap 15 to tap 30, the second-harmonic distortion builds until it is at its peak value about 1 octave before tap 30 (tap 20). After tap 20, however, it is gradually filtered away because the second-harmonic frequency begins to fall in the cutoff region of the cochlear filters. By the time that the signal is at tap 30, there is only a small amount of distortion left. Thus, the sharp cochlear rolloff ensures that each tap does not suffer much distortion at its BF.

Figure 26 illustrates that, at the BF, the output amplitude and harmonic distortion barely change with amplitude for amplitudes beyond about 40mV or 50mV. The second harmonic is 25dB smaller than the first harmonic for a wide range of input amplitudes. The reduction in harmonic distortion is due to the accumulated effects of the action of the AGC at each cochlear stage, and to the sharp cochlear rolloff. Note that, in the corresponding harmonic-distortion plots for a single cochlear stage (Figure 17(b)), the second harmonic distortion at BF is only down a factor of 8 at 1V rms, and there is continued growth of all harmonics with input amplitude.

Although the CF as measured by amplitude curves (Figure 13(b)), shifts, the CF_{90} as measured by phase curves (Figure 14) does not change appreciably. These findings for a single cochlear stage are echoed in the cochlea as well: At cochlear tap 30, as Figure 27 shows, the phase curves have a relatively invariant CF, although the gain curves (shown in Figure 23(b)) shift with amplitude. The kinks and rising parts of the phase curves of Figure 27 are due to parasitic capacitances in the cochlear filters.

3.4. Noise, Dynamic Range, and SNR

Figure 20 illustrates that the noise at the output of a cochlear tap has contributions from the input-referred noise of each cochlear filter preceding that tap. To evaluate the total noise at the output we need to evaluate the noise per unit bandwidth of each of these sources, to evaluate the transfer function from each source to the output, to accumulate the contributions from the various sources, and then to integrate over all frequencies.

If there are N_{oct} filters per octave, then the frequency ratio r between the τ of any filter and the τ of the filter just to the right of that one is given by

$$\begin{aligned} r &= 2^{-1/N_{oct}}, && (46)\\ &= e^{-1/(N_{oct}/\ln(2))}, && (47)\\ &= e^{-1/N_{nat}}. && (48)\end{aligned}$$

Thus if $x = \omega\tau$ is the normalized frequency corresponding to the output tap of interest, then the filtering effects of the filters preceding the output filter are represented by $H(x)$, $H(rx)$, $H(r^2x)$,...as shown in Figure 20. Similarly, if I were the bias current at the output cochlear tap (corresponding to $\sqrt{I_1 I_2}$ at each tap), the bias current of the k_{th} preceding filter would be given by I/r^k.

From Eqs. (12) and (17), the normalized input-referred noise per unit bandwidth is given by

$$v_0^2(x)dx = \left(\frac{NqV_L}{2\pi QC}\right)\left(\sqrt{C_R} + \frac{x^2}{\sqrt{C_R}}\right)dx. \quad (49)$$

It is then easy to show that the input-referred noise per unit bandwidth for the k_{th} filter preceding the output tap is given by

$$v_k^2(x)dx = \left(\frac{NqV_L r^k}{2\pi QC}\right)\left(\sqrt{C_R} + \frac{x^2 r^{2k}}{\sqrt{C_R}}\right)dx, \quad (50)$$

since $x \to r^k x$ and $I \to I/r^k$. If there are M preceding taps, then the total output noise per unit bandwidth $v_{out}^2(x)dx$ is given by

$$\frac{NqV_L}{2\pi QC}\sum_0^M r^n(\sqrt{C_R} + \frac{x^2 r^{2n}}{\sqrt{C_R}})\prod_0^n |H(r^k x)|^2 dx. \quad (51)$$

We obtain the total output noise at the cochlear tap of interest by integrating Eq. (51) over all x from 0 to ∞. Although the expression for the output noise at a cochlear tap can be written down, it is hard to solve in closed form. But it can be measured easily with a SR780 Spectrum Analyzer. Figure 28 shows what the noise spectrum of a cochlear tap looks like at tap 31 of our cochlea. It has a form predicted by Eq. (51) except for the second and third harmonic peaks; these peaks are due to nonlinearities in the filters.

Figure 29illustrates that the dynamic range at the output of tap 30 of our cochlea is greater than 60 dB at the BF of that tap (162Hz): Figure 29(a) shows the noise spectrum of the background noise at tap 30 which yields a total integrated noise of 50mV rms. When a BF sinusoidal signal (162Hz) of 0.907mV rms magnitude is applied to the input of the cochlea, it is amplified up by a factor of 57.1 to 51.8mV. Thus, the rms power of the signal and noise at tap 30 is about 72mV rms ($\sqrt{(51.8^2 + 50^2)}$). Now, at an output SNR ratio of 1, we would expect the signal and noise to have an rms power of $50\sqrt{2}$ = 70.7mV rms. The fact that the rms power is 72mV means that our minimum detectable signal, which corresponds to an output SNR of 1, is actually below 0.907mV. In fact, since the system is linear at small input amplitudes, the minimum detectable signal is 50mV/57.1 = 0.875mV. Figure 29(b) shows that the harmonic distortion at a 1V rms input is about $\sqrt{(11.85^2 + 2.82^2 + 1.11^2 + 0.245^2)}/315.8$ = 3.87%. This value is less than 4% which is commonly used as a measure of the upper limit of dynamic range of measuring-amplifier systems. Thus, at BF, we can process input signals over a ratio of $1000/0.875$ = 1143 in amplitude, or 1.306×10^6 in intensity. This range of intensity corresponds to a dynamic range of $10\log_{10}(1.306 \times 10^6)$ = **61.1dB**.

At large signals, the SNR at BF improves for two reasons: The signal amplitude gets larger—though not in a linear fashion, because of the AGC–and the noise amplitude drops, because of the lowering of Q. Figure 30(a) illustrates this effect for a 1V input and for a 0.9mV input. Figure 30(b) shows a plot of the signal amplitude and the noise amplitude for various input levels. The signal amplitude was evaluated as the square root of the power at the BF in spectral plots like those in Figure 30(a); the power at the harmonic peaks was ignored, although the power at these peaks also is due to the signal. We evaluated the noise power by integrating the power over all frequencies in the noise spectrum. The noise spectrum was obtained by removing all signal and harmonic peaks in the spectrum. We interpolated the noise spectrum in the regions where we removed the peaks. The noise amplitude was the square root of the noise power.

Figure 31 shows a plot of the SNR (signal power/noise power) as a function of input amplitude. As the input rms amplitude changes by a factor of about 61dB in intensity (0.9mV to 1V rms), the SNR changes by a factor of about 31dB (1 to 1241).

Figure 32 shows how our AGC cochlea extends the dynamic range of a hypothetical linear low-Q cochlea. The linear low-Q cochlea can be viewed as being representative of just the passive basilar membrane, with no outer hair cells [4]. Thus, we call our AGC cochlea with amplification (high-Q) an *active cochlea*, and the linear low-Q cochlea a *passive cochlea*. Some silicon cochleas have been built with a passive cochlea acting as a front end to a bank of bandpass filters [5].

Suppose that the passive cochlea has the same gain, and the same low Q, as the active cochlea at the largest input levels of 1V rms. Both cochleas will then have the same low-Q noise floor of 8.96mV at 1V. Since the passive cochlea maintains the same 0.315 gain at all intensities, its minimum detectable signal is given by 8.96mV/0.315 = 28.4mV. The active cochlea has a high Q at small input levels such that it amplifies the input signal and the noise. At BF, however, it amplifies the signal significantly more than the noise. In fact, its minimum detectable signal occurs when a 0.82mV input at BF has been amplified up by a factor of 59 to be just above the noise floor, which has now increased to 48.4mV.[2] Thus, the active cochlea extends the dynamic range of the passive cochlea by having the minimum detectable signal decrease by $20\log_{10}(28.4/0.82)$ dB = 31dB! It is known that outer hair cells in the biological cochlea extend the lower end of our dynamic range of hearing by about 40dB.

Figure 33(a) shows the noise spectra of various taps from tap 1 to tap 37. The corner frequency of the noise spectra successively decrease from tap 1 to tap 37, while the noise per unit bandwidth successively increases. The peak height increases and converges to an asymptotic limit as we traverse the cochlea. Note that, because of offsets in the cochlea, there is an abrupt reduction in corner frequency between taps 21 and 25. This abrupt reduction in bandwidth lowers the noise below the theoretical value that we would expect for taps close to and beyond this region. The total integrated noise over all frequencies is shown in Figure 33(b). The noise increases due to accumulation and amplification as we traverse the filters of the cochlea. However, the successive lowpass filtering limits this growth, until, in the asymptotic limit, there is an equilibrium between noise accumulation and noise filtering, and the noise ceases to grow. The discontinuities in the curve around tap 21 to tap 25 are due to the abrupt reductions in bandwidth around this region. The eventual convergence of the noise is due to the exponential taper of the cochlea: The exponential taper results in an accumulation of noise terms with coefficients that are determined by the terms of a geometric series with geometric ratio r (Eq. (51)). Since $r < 1$, the geometric series converges. Note that, as we increase the number of filters per octave, by Eq. (46), we increase r, and the noise increases. There is thus a tradeoff between the sharpness of the rolloff slope of the cochlea, which increases with N_{oct}, and noise reduction. The noise is also extremely sensitive to the value of Q because of the sensitive dependence of the $H(r^k x)$ terms of Eq. (51) on Q. We used Q = 1.5, and N_{oct} = 11.2 as a good compromise between not having too much noise at the output of the cochlear taps, and not having broad filters with shallow rolloff slopes.

Figure 34(a) shows the minimum detectable signal and maximum undistorted input at the BF of each tap in the cochlea. The minimum detectable signal was measured as described earlier in this section. The maximum undistorted input was measured by finding the input rms value at which the second harmonic (by far the dominant harmonic) was attenuated by 25 dB when compared with the first harmonic. We observe that the maximum undistorted input is nearly constant at 1V rms, except for the first few taps, where the action of the strong AGCs at each tap have not accumulated sufficiently to reduce the distortion.[3] Figure 34(b) shows the dynamic range at various taps. The dynamic range varies from about 59dB to 64dB. The early taps have little accumulation of noise or gain, in contrast with the late taps, which have large accumulation of noise and gain. The effect of gain regulation in the AGC causes the accumulation of noise and of gain to be approximately in balance, so the dynamic range does not suffer a huge variation across taps. However, as we would expect, Figure 35 shows that the maximum output SNR at BF (the SNR at 1V rms input) falls as we traverse the cochlea. It is maximum at tap1 (6.74×10^4 or 48.3dB) where there is the least noise, and minimum at tap 37 where there is the most noise (649 or 28.1dB). The discontinuities, due to the CF offset of the cochlea around taps 21 to 25, are evident in Figure 34 and Figure 35.

3.5. Spatial Characteristics

Figure 36(a) shows the spatial response of various taps of the cochlea to a 162Hz input, for various amplitudes. To understand the similarity of Figure 36(a) to Figure 23(a), we can view the cochlea as performing a frequency-to-place transformation with $-log(f) \rightarrow x$ [4]. Even the harmonic-distortion plot of Figure 36(b) is quite similar to that of Figure 25(a). The most severe distortion occurs at a place that corresponds to a corner frequency that is 1 octave higher than the corner frequency at the best place (BP). Figure 37 shows the shift in BP for two different frequency inputs to the cochlea.

3.6. Dynamics of Gain and Offset Adaptation

Figure 38(a) shows the attack response of cochlear tap 30 to the abrupt onset of a tone at the tap's BF (162Hz). After a transient at the first cycle, the envelope of the response adapts quickly to the new intensity, corresponding to the quick onset adaptation of the peak detector. The offset correction has a slower rate of adaptation and continues to adapt with some ringing even after the envelope adaptation is complete.

Figure 38(b) shows the release response of cochlear tap 30 to an abrupt decrease in the intensity of the BF tone. The adaptation of the envelope is much slower than that shown in Figure 38(a) because of the slow adaptation of the peak detector to inputs of decreasing intensity. The DC offset adaptation continues to have a rate of adaptation that is slower than the rate of envelope adaptation.

3.7. The Mid-Frequency and High-Frequency Cochleas

So far, we have dwelled almost entirely on the properties of the low-frequency cochlea; the properties of the other cochleas are similar. Figure 39 shows the variation in Q versus corner frequency due to bias-current differences in a cochlear filter. There is a variation in Q as we go from subthreshold behavior at low frequencies to above-threshold behavior at high frequencies. However, our high-frequency circuits operate in moderate inversion (near the graded transition from subthreshold to above threshold), and thus the change in Q is not significant. Figure 40 shows that, consequently the "sounds of silence", that is, the noise spectra at the various taps in the low, mid, and high-frequency cochleas are similar in shape across the entire frequency range (100Hz to 10kHz).

4. Analog Versus Digital

The total resting current consumption of all three of our cochlear cascades was measured to be 95μA. Playing microphone speech through our cochleas increased the power consumption to about 99μA. Thus, the total power consumption of our cochlea is about 100μA $\times$5V = **0.5mW**. Our area consumption was 1.6mm$\times$1.6mm$\times$3 = **7.7** mm^2 in a 1.2μm process. The pitch of a single cochlear stage, including all scanning circuitry and with a conservatively large number of power buses (to prevent unwanted coupling through the supplies), was 102 μm$\times$ 444 μm.

The high-frequency cochlea consumes more than 3/4 of this power. We can easily cut our power dissipation to 0.2mW by having a tilt on the V_A voltages, although we did not implement this tilt on our current design. If only telephone bandwidth is required, we can do away with the high-frequency cochlea and cut our power dissipation to 0.125mW. If we implement the tilt on the V_A voltages and do not use the high-frequency cochlea, then our power consumption reduces to 50μW.

We next compare the power and area consumption of our analog cochlea, an ASIC digital cochlea, and a noncustom microprocessor (μP) cochlea. We begin by describing the design of the ASIC digital cochlea.

4.1. The ASIC Digital Cochlea

Figure 41 shows a block-level schematic of a digital cochlea, similar to our analog cochlea, and described in [14]. Second-order recursive digital filters with tapering filter coefficients model the basilar membrane. Half-wave rectification circuits (HWR) perform MSB lookup to model the inner hair cells. Automatic-gain-control circuits (AGC) with cross talk model the olivocochlear efferent system. The multiscale AGC is modeled over 4 time scales.

This is a custom cochlea, designed to be as efficient in power and area consumption as possible. A digital input, clocked at 50 Khz, forms the input to the cochlea; that frequency is slightly over the Nyquist frequency of 36khz for the highest-frequency location of the cochlea, and is necessary to obtain robust behavior with the filtering and nonlinear operations in the cochlea. It is possible to implement a multirate sampling system, but calculations show that the bandwidth needed to implement 95 stages of the cochlea from 18Khz to 100Hz (as in the analog cochlea) is equivalent to the bandwidth needed to implement 17 stages at 18kHz. Thus, a multirate system can help only by a factor of 5.6. If the overhead in circuitry and complexity needed for a multirate system is factored in, there may be no advantage whatsoever. Thus, we shall confine ourselves to a system with only one rate of sampling. Note that we need only 95 stages in the digital cochlea (as opposed to 117 stages), since we do not need the redundancy of the overlapping-cascades architecture. To handle the input dynamic range of 60dB, (i.e., 10 bits), it is necessary to do fixed-point operations at a precision of approximately 24 bits; otherwise, overflow errors and round-off–error accumulation can seriously jeopardize the computation.

The system shown in Figure 41 is implemented most efficiently with a bit-serial representation, where the bits are processed serially, and each filter, HWR, and AGC block is reused 95 times to compute the effect of the entire cascade. The reuse of circuitry results in tremendous savings in area and power, and makes a digital cochlear implementation feasible on a single chip. There is, of course, overhead in the storage that is necessary to implement these computations.

Table 1. Cochleas

	ANALOG	ASIC DIGITAL	DEC α
TECH.	1.2μm	0.5 μm	0.5μm
V_{DD}	5V	2V	3.3V
POWER	0.5mW	150mW	50W
AREA	7.7mm^2	25 mm^2	299 mm^2

The proposed ASIC digital cochlea was never built. However, we can estimate what its power dissipation would have been. The Clock Rate is 50 kHz $\times$ 95 stages $\times$ 24 bits = 114.0Mhz. The power supply would need to be about 2.0 V to attain a 114.0MHz clock rate. Let's assume that the technology is 0.5 μm. The number of gates needed for the computation is roughly 40 (number of gates for 1 multiply operation, including storage overhead) $\times$ 24 (number of bits) $\times$ 7 (3 multiplies in filter and 4 in the AGC) = 6720 gates + RAM and ROM. The 13 add operations comprising 5 adds in the filters and 4 $\times$ 2 adds in the AGC are treated as being essentially free in fixed-point computations. The gate.Hz = 6720 $\times$ x 114Mhz = 0.77×10^{12} gate Hz. The gate capacitance = (0.5 μm $\times$ 0.5 μm $\times$ 10 (transistors per gate) $\times$ 2 fF (cap. per unit area) = 50 fF. The switching energy per gate = 50 fF x $(2.0)^2 = 2.0 \times 10^{-13}$J. The power dissipation is therefore 0.77×10^{12} gate.Hz $\times$ 2.0×10^{-13} = 0.154W, which we shall round down to **0.15W**. The area we would need to build this chip is estimated to be 5 mm $\times$ 5 mm (in 0.5 μm tech.) = **25** mm^2.

4.2. μP cochlea

In FLOPS, we need about 50 Khz (bandwidth) $\times$ 95 (number of stages) $\times$ 20 (7 multiplies and 13 adds) = 95 MFLOPs to implement our cochlea. Note that adds cannot be treated as free in floating-point operations. On the specfP92 Ear program, the DEC 21164 running on an Alpha server 8200 5/300 does about 1275 times better than a Vax 11/780. The Vax 11/780 is specified at 0.1 MFLOPS. Thus, the DEC α is capable of 1275 $\times$ 0.1 = 127.5 MFLOPS which is enough for our computation, The DEC α consumes **50 W** and has an area of 16.5 mm $\times$ 18.1 mm = **299** mm^2.

4.3. Comparison of Analog and Digital Cochleas

Table 1 compares the power and area consumption of the various cochleas. Note that our analog cochlea would be more efficient in area by about a factor of 2 to 4 if it were also implemented in a 0.5μm technology like the digital designs. However, we have not scaled down the analog numbers; we have just shown them for our current 1.2μm technology.

The analog implementations are more efficient in power than are custom digital implementations by a factor of 300, and than are noncustom μP implementations by a factor of 1×10^5. The analog cochlea can run on 1Ah batteries for more than a year (with 100μA current consumption), whereas the best digital cochlea would be able to run for only less than 1 day (with 75mA current consumption).

The area comparisons show that, even in an inferior technology (1.2 μm vs. 0.5 μm), the analog cochlea is about 3 times more efficient than is the custom ASIC cochlea, and is about 40 times more efficient than is the microprocessor implementation.

The cochlear comparisons were generous to digital implementations: We used a better technology (0.5 μm versus 1.2 μm), operated with a power-saving supply voltage (2.0 V versus 5.0 V), used an efficient bit-serial implementation, did not include the cost of the 10-bit or 13-bit A/D converter, and were more conservative in our cost estimates. Nevertheless, the analog implementations were two to five orders of magnitude more efficient than the digital implementations. To compete with digital systems, the analog systems had to be designed with wide-dynamic-range circuitry, and had to compensate for their offsets. In fact, most of the analog cochlea's resources in area were expended in filter linearization, low-noise transduction, and offset-compensation circuitry. Most of the analog cochlea's resources in power were expended in low-noise sensing circuitry. The number of devices needed to do the actual computation was nevertheless so small that 117 stages could be implemented easily on one chip, with room to spare.

By contrast, the digital cochlea's resources in area and power were not primarily consumed in maintaining precision, although extra bits were necessary to prevent overflow and roundoff errors. Rather, the actual computation was so expensive in digital that only one stage of the cochlear cascade was feasible on a single chip. That stage had to be reused 95 times in succession, at a fast rate of 114MHz, to finish the computation in real time. In other words, the analog implementation was slow per computational stage, cheap, and completely parallel. The digital implementation was fast per computational stage, expensive, and fully serial. We might wonder—if the digital implementation were slow and fully parallel just like the analog one, would the comparisons in efficiency seem less drastic? The answer is yes for power consumption because it could be reduced by turning down the power-supply voltage and clock frequency. The answer is no for area consumption, because it would be 95 times worse. In this particular case, however, the size of the chip required for the parallel digital implementation would be totally unfeasible. In other words, there is no free lunch: the inefficiency of using a transistor as a switch will always show up somewhere.

5. The Biological Cochlea

The biological cochlea is far more complex than is our electronic cochlea, and it is surprising that we can replicate much of its functionality with just our simple circuits. Our aim is not to replicate its functions exactly, as computer modeling attempts to do, but rather to exploit its clever computational ideas to build more efficient electronic architectures for artificial hearing. Such architectures may enable the design of superior hearing aids, cochlear implants, or speech-recognition front ends. In addition, as we shall show in Section 5.1, the synthesis of an artificial cochlea can help us to improve our understanding of how the biological cochlea works.

The functions of the biological cochlea that we can replicate are:

1. The frequency-to-place transformation, as implemented by the amplification and propagation of traveling waves
2. A compressive nonlinearity at and beyond the BF of a cochlear tap. Like the biological cochlea, our response is linear for frequencies well below the BF. Our compression is achieved through an AGC. In the biological cochlea, it is still a matter of debate as to how much of the compression arises from a dynamic AGC and how much from a static nonlinearity. We have reported on cochleas where the compression arises solely from a static nonlinearity as well [9].
3. An asymmetric attack and release response to transient inputs.
4. The extension of dynamic range due to active amplification. Our dynamic range is extended from 30dB to about 60dB. In the biological cochlea, it is believed that amplification by outer hair cells extends the dynamic range of the cochlea by about 40dB.

5. The broadening of the pattern of excitation as the input intensity is increased. The dual effect, which we can also model, is the broadening of the frequency-response curves as the input intensity is increased.
6. The shift of the peak frequency towards lower frequencies as the input intensity is increased. The dual effect, which we can also model, is the shift of the peak place of excitation toward the input of the cochlea as the intensity is increased.
7. A sharp cochlear roll-off slope.
8. Masking of adjacent frequencies and harmonics due to the effects of the AGC and nonlinearity, respectively. However, our dominant harmonic is the second harmonic. In the biological cochlea, the dominant harmonic is the third harmonic.

5.1. Traveling-Wave Architectures Versus Bandpass Filters

Why did nature choose a traveling-wave architecture that is well modeled by a filter cascade instead of a bank of bandpass filters? We suggest that nature chose wisely, for the following three reasons:

1. To adapt to input intensities over a 120dB dynamic range, a filter bank would require a tremendous change in the Q of each filter. To compress 120dB in input intensity to about 40dB in output intensity the filter Qs must change by 80dB; a dynamic-range problem in the input is merely transformed into a dynamic-range problem in a parameter. In contrast, in a filter cascade, due to the exponential nature of gain accumulation, enormous changes in the overall gain for an input can be accomplished by small distributed changes in the Q of several filters.
2. Large changes in the Q of a filter are accompanied by large changes in the filter's window of temporal integration. Thus, in filter banks, faint inputs would be sensed with poor temporal resolution, and loud inputs would be sensed with good temporal resolution. In contrast, in a filter cascade, the shifts in temporal resolution with intensity change only in a logarithmic fashion with intensity, as opposed to in a linear fashion as in the filter bank.
3. A sharp rolloff slope in a filter is extremely useful in limiting distortion, and in enhancing spectral contrasts. A sharp rolloff slope arises naturally in the cochlear filter cascade. To accomplish such a rolloff slope in a filter bank requires very high-order filters, and consequently an enormous amount of circuitry at each tap. In contrast, in the filter cascade, the burden of creating a high-order rolloff is shared collectively, so only one new filter needs to be added for each new desired corner frequency.

There are two problems that need to be addressed in a filter cascade:

1. A filter cascade is prone to noise accumulation and amplification. The solution to this problem is either to have an exponential taper in the filter time constants such that the output noise converges (the solution found at high CFs in the biological cochlea), or to limit the length of the cascade (the solution at low CFs in the biological cochlea). The exponential taper also results in elegant scale-invariant properties.
2. The overall gain is quite sensitive to the value of each filter's Q. The solution to this problem is to have gain control regulate the value of the Q's in the cascade. If the gain control is sufficiently strong, then the collective adaptation in Q across many filters will compress a wide input dynamic range into a narrow output dynamic range.

6. Applications to Cochlear Implants

Front-end modules in current cochlear implant devices make use of parallel banks of independent bandpass filters. For example, the front-end module of a state-of-the-art commercial multichannel cochlear implant devices consists of 20 fourth-order bandpass filters with center frequencies between 250Hz and 10kHz. The filters are implemented using switched-capacitor techniques. The total power dissipation of such implementations is on the order of several milliwatts, and the dynamic range is only 35 to 40 dB.

Our neuromorphic approach mimics several aspects of the biological cochlea, as described in Section 5. In addition, our dynamic range exceeds 60dB. Our power dissipation for a 117-stage cochlea with a roll-off slope corresponding to a high-order filter (10th order to 16th order) is 0.5mW. If we use fewer stages and fewer filters per octave to correspond to current values in implant front ends, we could, we estimate, cut our power dissipation to 50 μW. This power dissipation is about 20–100 times lower than that in current front ends.

Thus, in terms of biological realism, dynamic range, and power we can do much better than current implant front ends. Previously [11], we described how a nonlinear center-surround operation on the outputs of the cochlear taps can convert cochlear lowpass information into bandpass information without degrading the temporal resolution at that tap. A neuromorphic front-end module like ours satisfies the fundamental requirements of future cochlear-implant speech processors [15].

7. Conclusions

We described a 117-stage 100Hz-to-10kHz cochlea that attained a dynamic range of 61dB while dissipating 0.5mW of power. The wide dynamic range was attained through the use of a wide-linear-range transconductance amplifier, of a low-noise filter topology, of dynamic gain control (AGC), and of an overlapping-cascades architecture. An infrastructure of automatic offset adaptation, small amounts of low-frequency attenuation, and scale-invariant BiCMOS circuit techniques provided robust operation. The low power, wide dynamic range, and biological realism suit our cochlea to be used as a front end for cochlear implants. The design of our electronic cochlea suggests why nature preferred an active traveling-wave mechanism over a bank of bandpass filters as a front end for hearing.

Notes

1. We are assuming that the supply voltage limits the range of operation of the system. If there is some other voltage that limits the range of operation of the system, then power is wasted through an unnecessarily high supply voltage. We choose not to operate the system in this nonoptimal situation.
2. These numbers (gain of 59, noise of 48.4mV, and minimum detectable signal of 0.82mV) are slightly different from the numbers that we quoted earlier (gain of 57.1, noise of 50mV, and minimum detectable signal of 0.875mV) because of the interpolation procedures used in our data processing algorithm, and because of the different times at which the data were collected.
3. We were able even to apply a 1.4V input rms signal and to keep the distortion under 25 dB (due to the strong AGC), but we refrained from doing so because the input signal then would be just at the edge of our DC operating range; operating the cochlea at this extreme is possible, but we chose not to so as to leave a safety margin.

References

1. R. Sarpeshkar, T. Delbrück, and C. Mead, "White Noise in MOS Transistors and Resistors", *IEEE Circuits and Devices*, Vol. 9, No. 6, pp. 23–29, Nov. 1993.
2. B. M. Johnstone, "Genesis of the Cochlear Endolymphatic Potential", *Current Topics in Bioenergetics*, Vol. 2, pp. 335–352, 1967.
3. B. C. J. Moore, *An Introduction to the Psychology of Hearing", Third Edition*, pp. 47–83, Academic Press Limited, London, 1989.
4. C.A. Mead, *Analog VLSI and Neural Systems*, pp. 179–192, 279–302, Addison-Wesley Publishing Co., Reading, MA, 1989.
5. W. Liu, *An Analog Cochlear Model: Signal Representation and VLSI Realization*, PhD. Thesis, John Hopkins University, Baltimore, Maryland, 1992.
6. L. Watts, D. A. Kerns, R. F. Lyon, and C.A. Mead, "Improved Implementation of the silicon cochlea", *IEEE Journal of Solid-State Circuits*, 27(5), pp. 692-700, May 1992.
7. N. Bhadambkar, "A variable resolution nonlinear silicon cochlea", Technical Report CSL-TR-93-558, Stanford University, 1993.
8. A. van Schaik, E. Fragniere, and E. Vittoz, "Improved Silicon Cochlea using Compatible Lateral Bipolar Transistors", *Advances in Neural Information Processing Systems 8*, pp. 671–677, edited by D. Touretzky, et al., MIT Press, Cambridge, MA 1996.
9. R. Sarpeshkar, R. F. Lyon, and C. A. Mead, "An Analog VLSI Cochlea with New Transconductance Amplifiers and Nonlinear Gain Control", *Proceedings of the IEEE Symposium on Circuits and Systems*, Atlanta, Vol. 3, pp. 292–295, May 1996.
10. R. Sarpeshkar, R. F. Lyon, and C. A. Mead, "A low-power wide-linear-range transconductance amplifier", *Analog Integrated Circuits and Signal Processing*, Vol. 13, No. 1/2, pp. 123–151, May/June 1997.
11. R. J. W. Wang, R. Sarpeshkar, M. Jabri, and C. Mead, "A Low-Power Analog Front-end Module for Cochlear Implants", presented at the XVI World Congress on Otorhinolaryngology, Sydney March 1997.
12. M. A. Ruggero, "Response to Sound of the Basilar Membrane of the Mammalian Cochlea", *Current Opinion in Neurobiology*, Vol. 2, pp. 449–456, 1992.
13. T. D. Clack, J. Erdreich, and R. W. Knighton, "Aural Harmonics: The Monoaural Phase Effects at 1500Hz, 2000Hz, and 2500Hz Observed in Tone-on-Tone Masking When f_1 = 1000Hz", Journal of the Acoustical Society of America, Vol. 52, Number 2 (Part 2), pp. 536–541, 1972.
14. C. D. Summerfield and R.F. Lyon, "ASIC Implementation of the Lyon Cochlea Model", *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, San Fransisco, March 1992.
15. G. Clark, "Cochlear Implants: Future Research Directions", *Annals of Otology, Rhinology, and Laryngology*, Vol. 104, No. 9, Part 2, pp. 22–27, 1995.

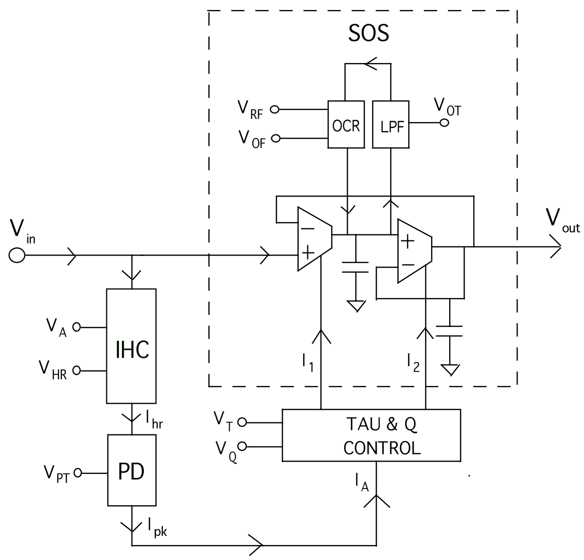

Fig. 1. Schematic for a Cochlear Stage. A single cochlear stage is composed of a filter (SOS) with offset-adaptation circuitry (LPF and OCR), an inner-hair-cell and peak-detector circuit (IHC and PD), and a tau-and-Q control circuit.

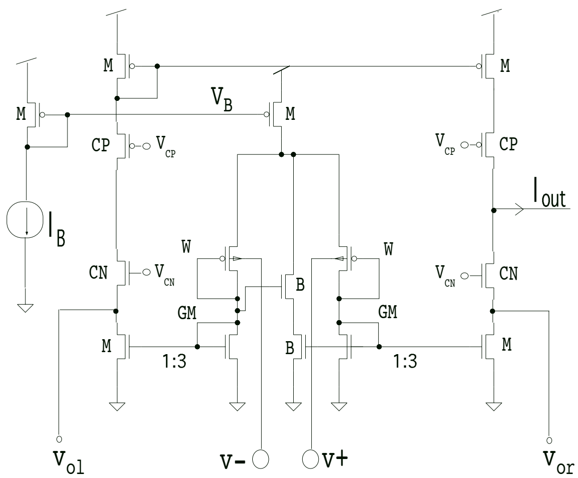

Fig. 2. The Wide-Linear-Range Transconductance Amplifier. The inputs to the amplifier are v_+ and v_-, and the output is the current I_{out}. The bias current is I_B. The voltages v_{ol} and v_{or} represent inputs from the offset-adaptation circuitry.

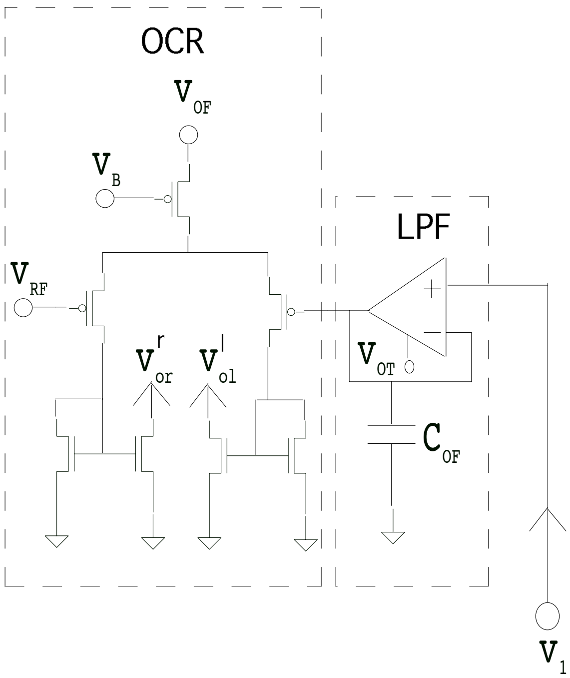

Fig. 3. The Offset-Adaptation Circuit. The input to the circuit is the output of the first amplifier of the SOS, V_1. The outputs v^1_{ol} and v^1_{or} connect to the corresponding inputs of the first amplifier of the SOS.

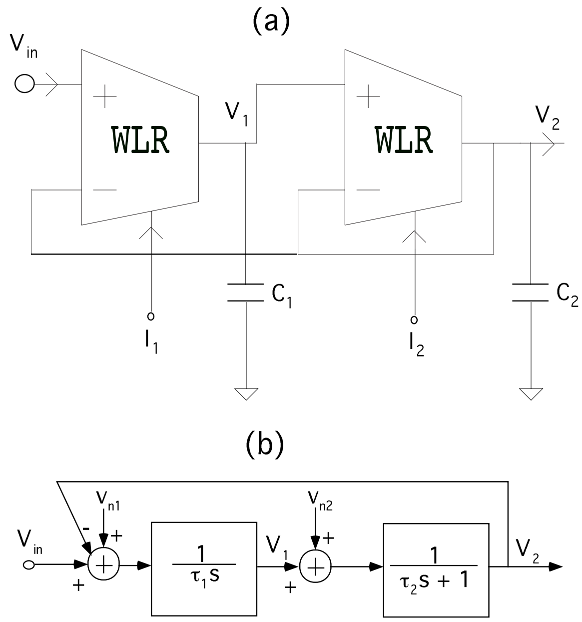

Fig. 4. The Second-Order Filter Circuit. (a) The input is V_{in}, the output is V_2, and the bias currents of the first and second amplifiers are I_1 and I_2. (b) The block-diagram equivalent of the filter is useful for making noise calculations. The voltages v_{n1} and v_{n2} represent the input-referred noise sources of the first and second amplifier respectively.

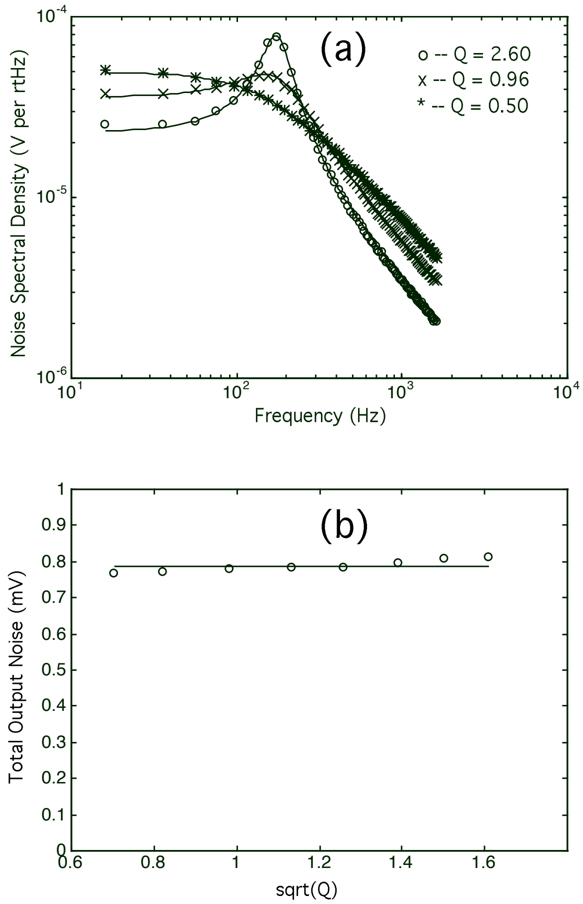

Fig. 5. Noise in the Second-Order Filter Circuit. (a) The noise spectrum changes shape as the Q of the filter is changed. (b) The total output noise integrated over all frequencies is approximately invariant with Q for this filter.

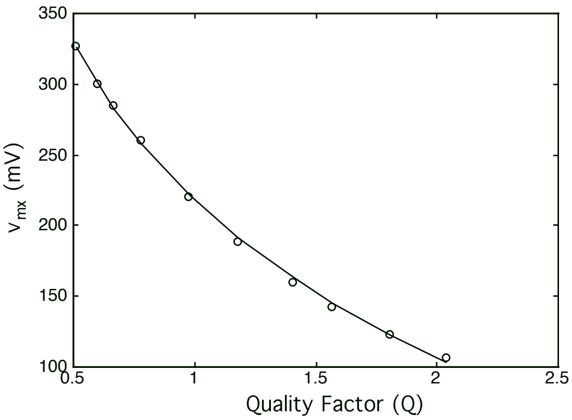

Fig. 6. Maximum Undistorted Signal in the Filter. The input amplitude at which the total harmonic distortion at the output is attenuated by 25dB with respect to the fundamental is plotted versus Q. The fundamental frequency is at the BF of the filter. The line is an empirical fit.

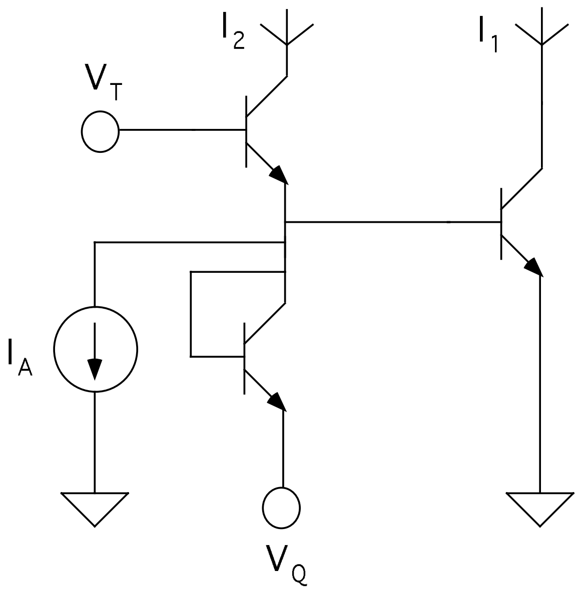

Fig. 7. Translinear tau-and-Q Biasing Circuit. The voltage V_T sets the τ of the filter, and the voltage V_Q sets the small-signal Q. The current I_A is a placeholder for a gain-control–correction current. The currents I_1 and I_2 are the bias currents of the first and second amplifiers of the filter.

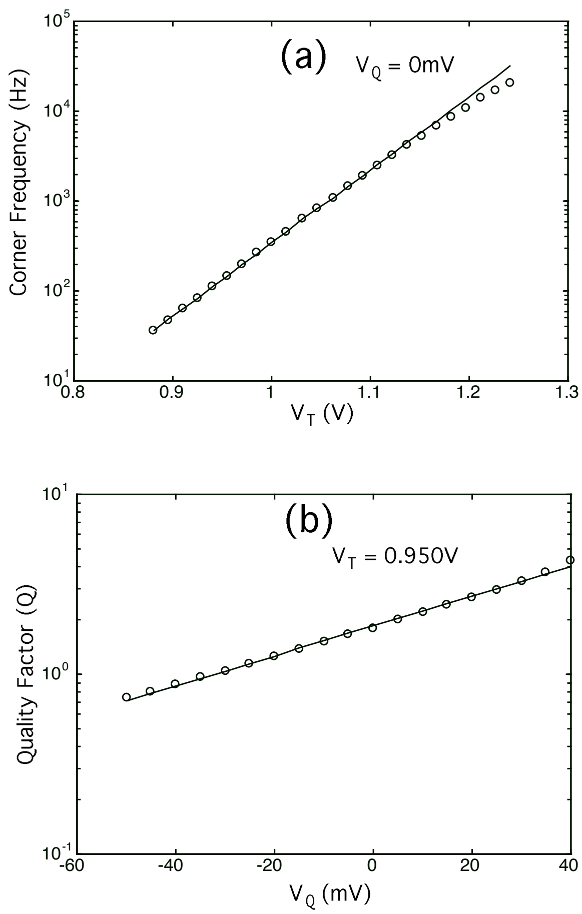

Fig. 8. Tau-and-Q Control Circuit Characteristics. (a) The corner frequency has an exponential dependence on the voltage V_T. (b) The quality factor Q has an exponential dependence on the voltage V_Q.

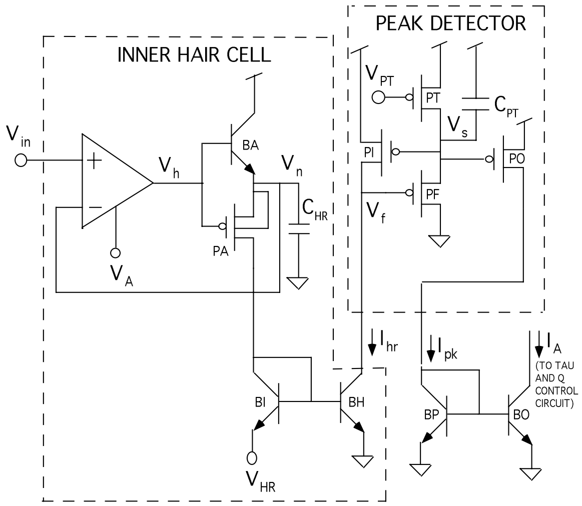

Fig. 9. The IHC and PD Circuits. The inner hair cell transduces its input V_{in} to a current I_{hr} that is then fed to the peak detector. The output of the peak detector I_{pk} is mirrored to the tau-and-Q control circuit as a gain-control–correction current.

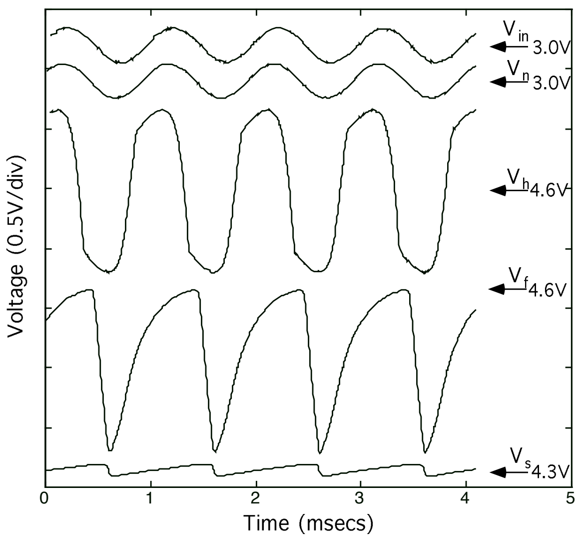

Fig. 10. The IHC and PD Circuit Waveforms. The waveforms for the voltages V_{in}–V_s illustrate the operation of the circuits of Figure 9.

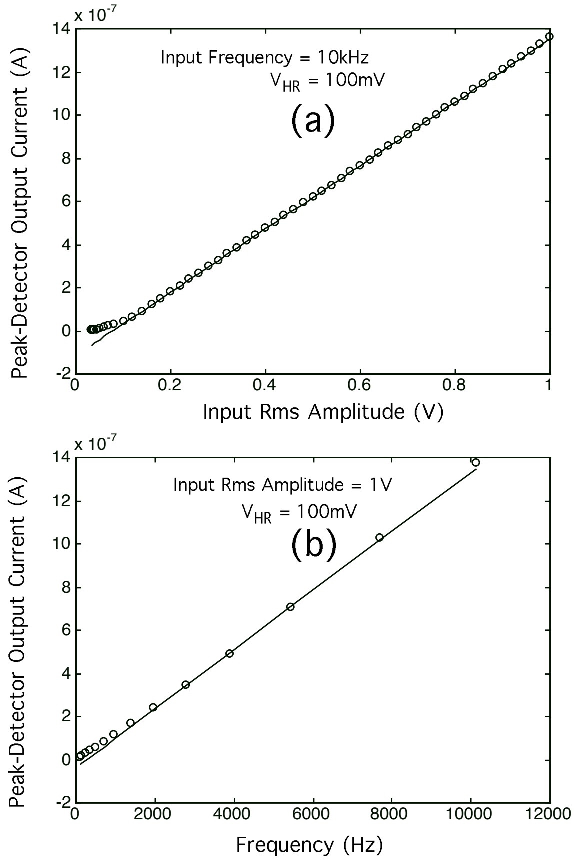

Fig. 11. IHC and PD Amplitude and Frequency Characteristics. (a) The current I_{pk} has a linear dependence on the input rms amplitude. (b) The current I_{pk} has a linear dependence on the the input frequency.

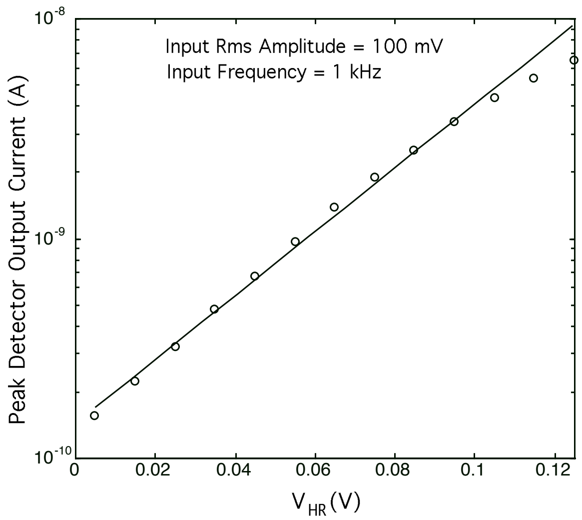

Fig. 12. Dependence of I_{pk} on V_{HR}. The current I_{pk} has an exponential dependence on the voltage V_{HR}.

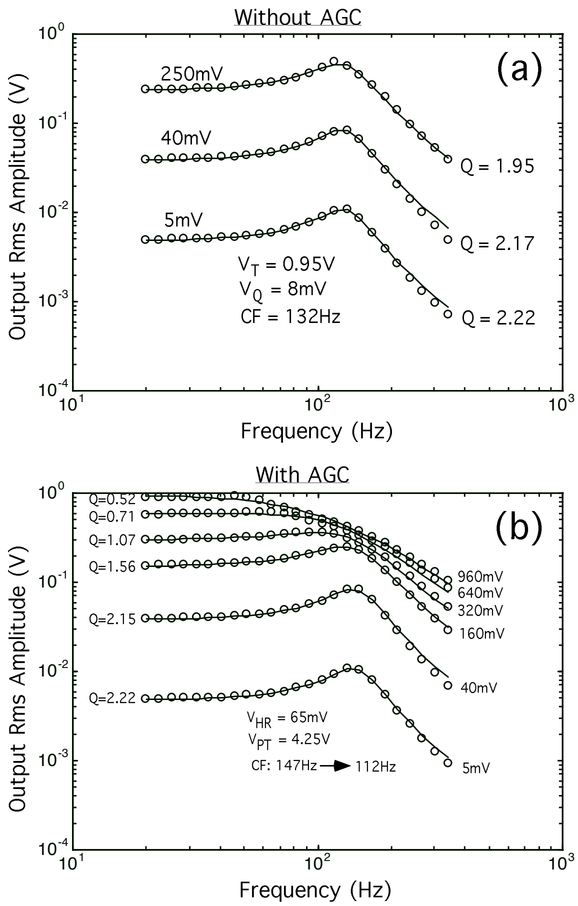

Fig. 13. Frequency-Response Characteristics of a Stage. (a) Without an AGC, it is impossible to obtain smooth and continuous data beyond an input rms amplitude of 250mV. (b) With an AGC, it is easy to obtain smooth and continuous data up to and beyond a 960mV rms input amplitude.

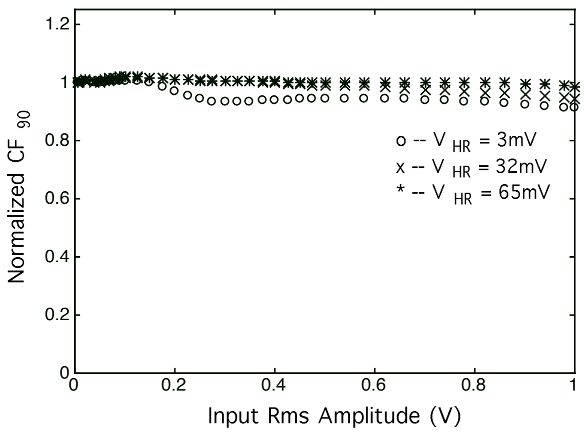

Fig. 14. CF_{90} Characteristics. The frequency at which the phase lag of the filter is 90 degrees is relatively invariant with input rms amplitude.

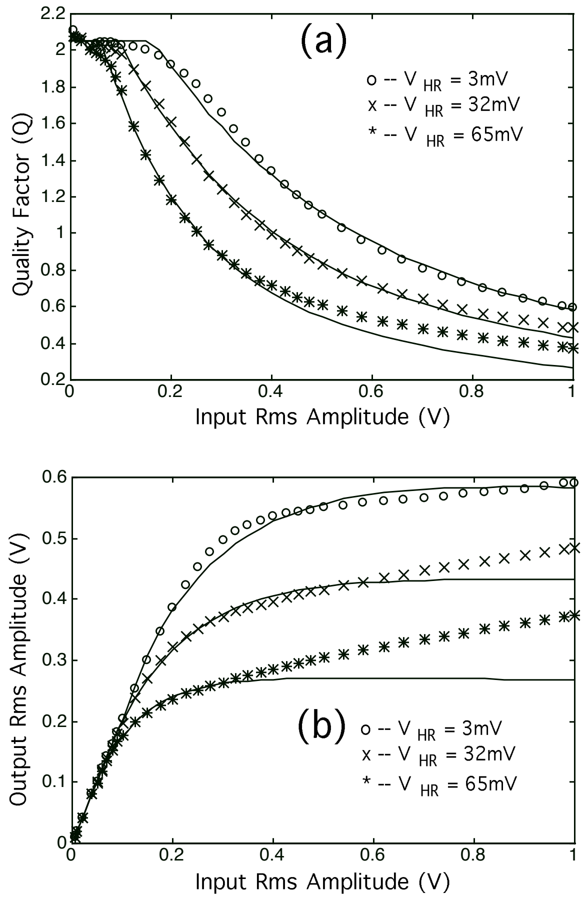

Fig. 15. Q-Adaptation Characteristics. (a) The Q adaptation due to the AGC is well fit by theory, except at large input rms amplitudes, and for strong AGC corrections (large V_{HR}). (b) The same data as in (a) except that we plot the output rms amplitude, instead of the Q.

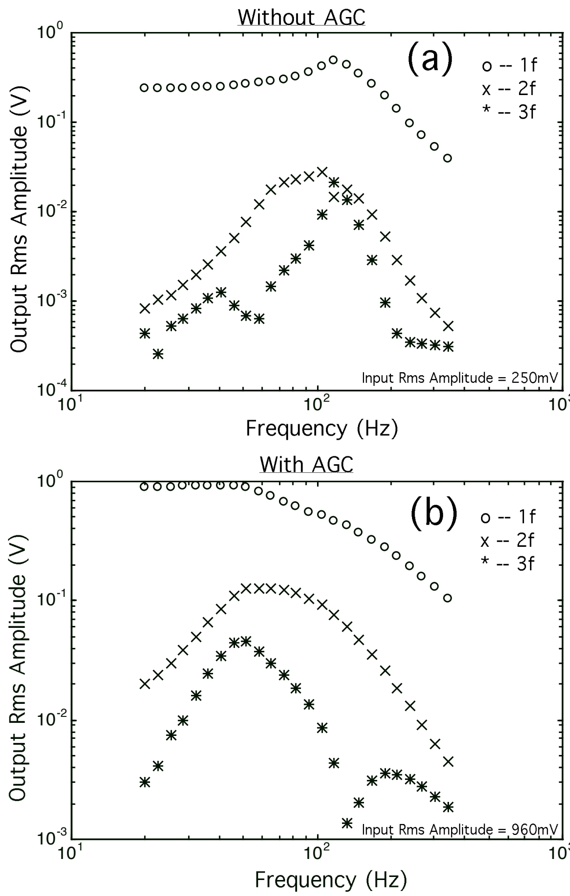

Fig. 16. Distortion Characteristics of the Filter. (a) Without an AGC, the distortion is already fairly high at a 250mV input rms amplitude. (b) With a strong AGC (V_{HR} = 65mV), the distortion is comparable to that in (a) at only a 1V rms input amplitude.

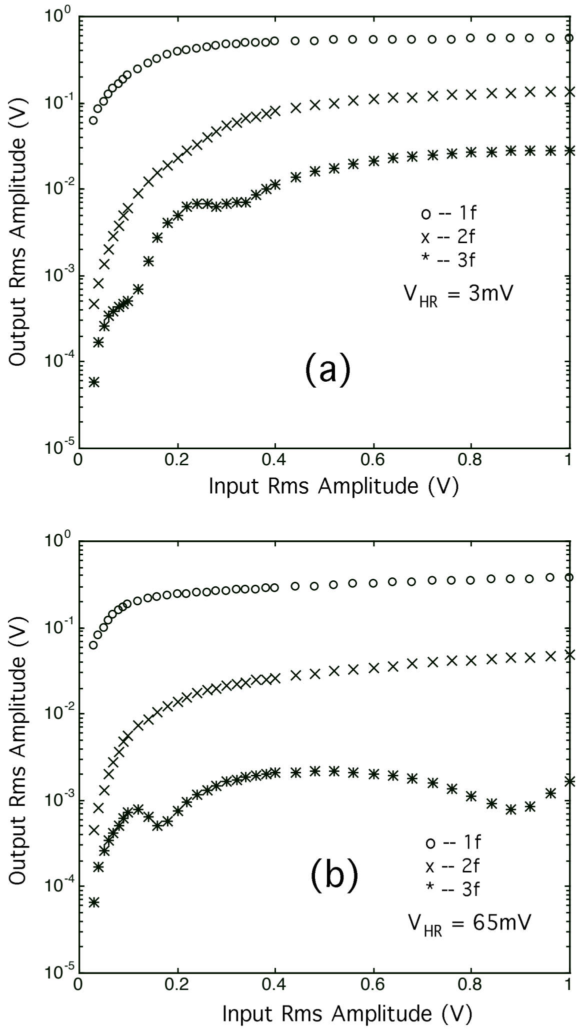

Fig. 17. Distortion Characteristics at BF. (a) With a weak AGC (V_{HR} = 3mV), the distortion levels are significant at a 250mV input rms amplitude. (b) With a strong AGC (V_{HR} - 65mV), the distortion levels are smaller than are those in (a) even at a 1V input rms amplitude.

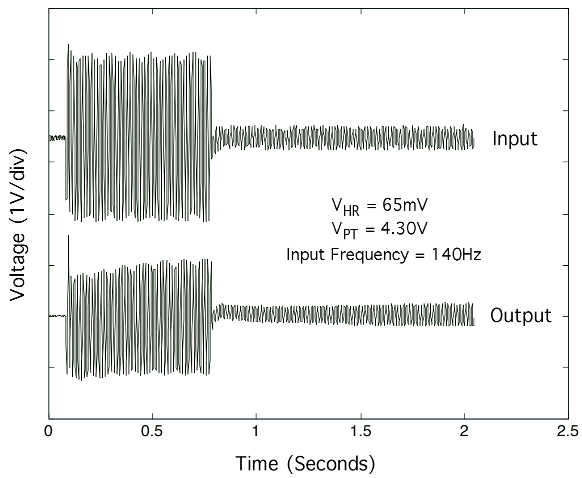

Fig. 18. Dynamics of Q-adaptation. The onset of a loud tone at BF preceded by a period of silence causes a brief transient on the first cycle, followed by a restoration to equilibrium. The reduction in intensity of the same tone from a loud to a soft value causes a gradual buildup in the output response as the gain of the AGC adapts.

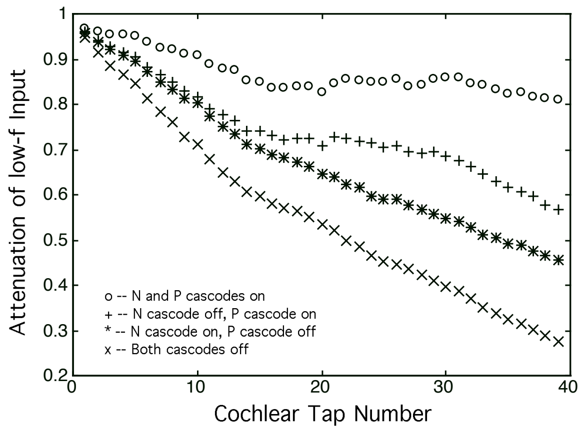

Fig. 19. Low-Frequency Attenuation in the Cochlea. The low-frequency attenuation for various conditions of open-loop amplifier gain are shown.

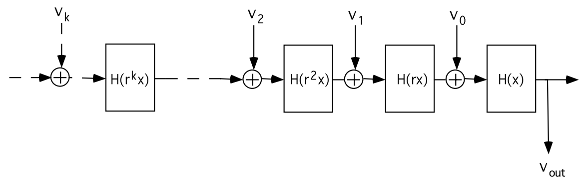

Fig. 20. Noise Accumulation in the Cochlea. The noise at the output tap of a cochlea, v_{out} is due to the accumulation, amplification, and filtering of noise from taps preceding that tap.

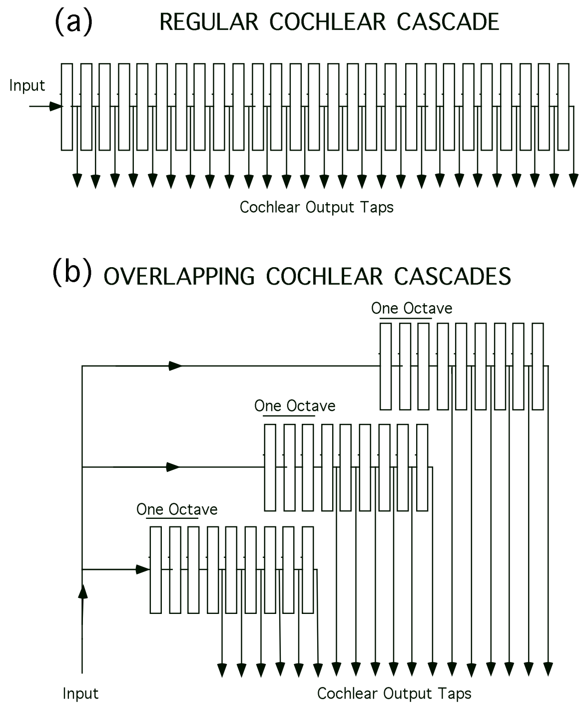

Fig. 21. Architecture of Overlapping Cascades. (a) In a regular cochlear cascade, the input is fed serially to all stages. (b) In an overlapping cochlear cascade, the input is fed in parallel to tiny cochlear cascades whose corner frequencies overlap by 1 octave.

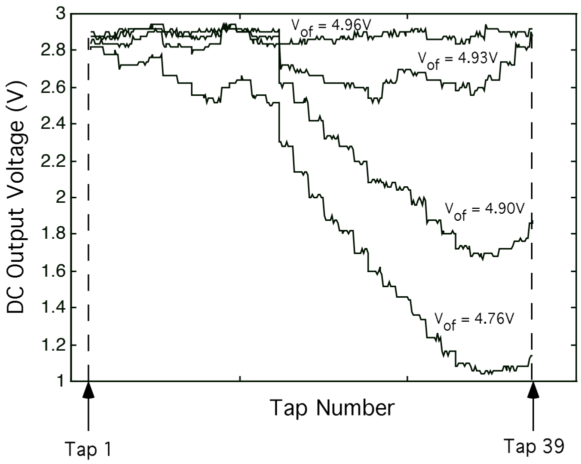

Fig. 22. Offset Adaptation in the Cochlea. As the loop gain of the offset-adaptation loop, controlled by V_{OF}, is increased, the offset accumulation across the taps of the cochlea is reduced.

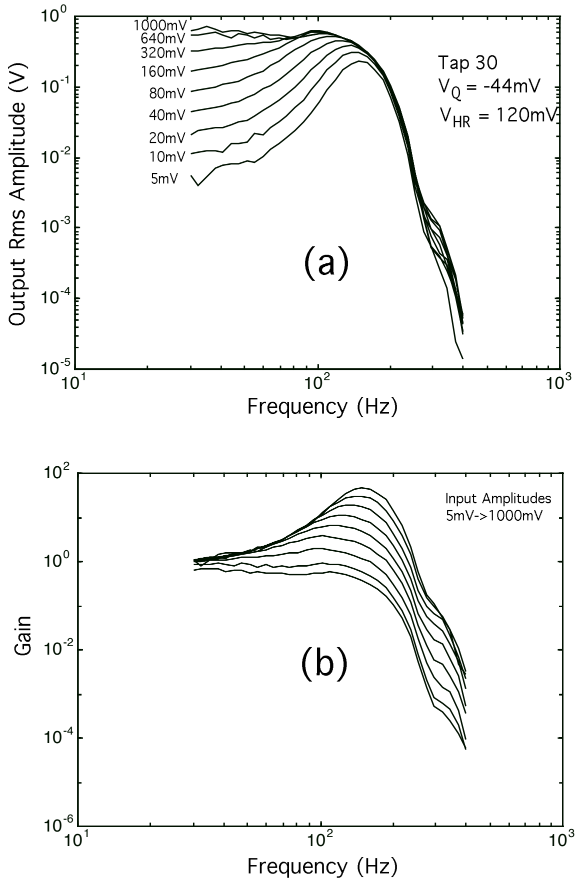

Fig. 23. Frequency-Response Curves of the Cochlea. (a) The frequency response for various input rms amplitudes is shown. (b) The same data as in (a) except that we plot the gain, instead of the ouput rms amplitude.

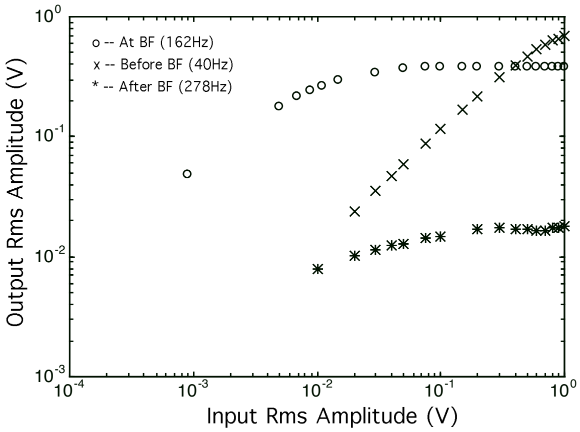

Fig. 24. Compression Characteristics of the Cochlea. The compression at a cochlear tap occurs primarily at and beyond the BF, whereas the response at frequencies below the BF is linear.

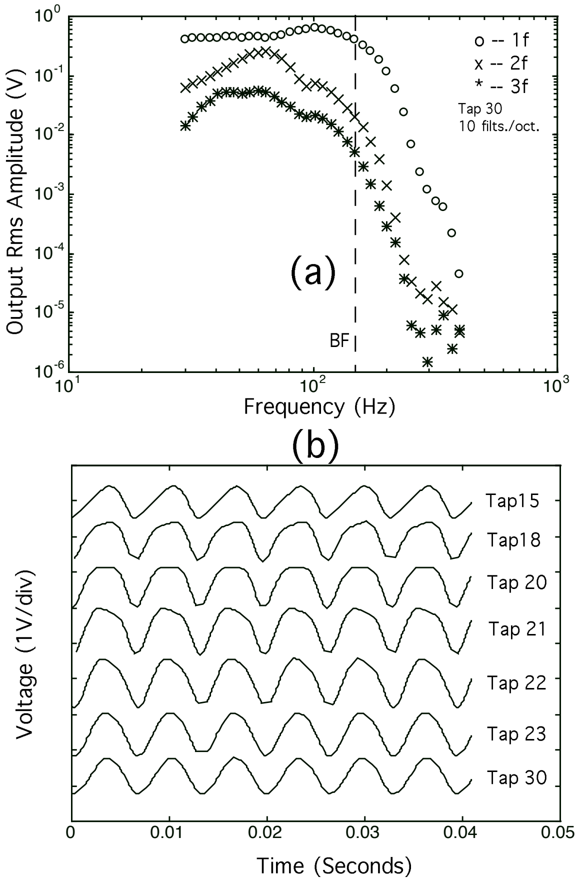

Fig. 25. Harmonic Distortion Characteristics of the Cochlea. (a) The harmonic distortion is most pronounced 1 octave before the BF, but is sharply attenuated at the BF. (b) The dual effect in space reveals that harmonic distortion is most pronounced 1 octave before tap 30 (at tap 20), but is filtered away by tap 30.

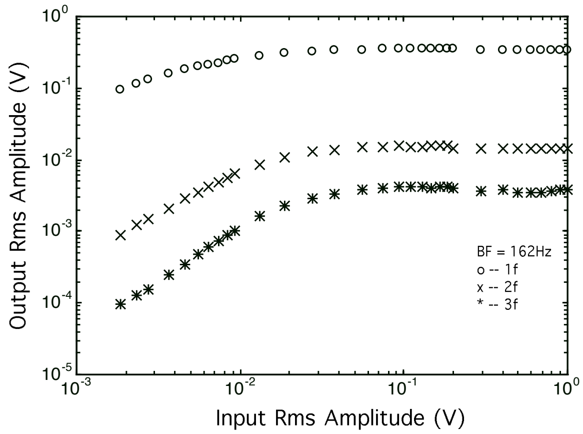

Fig. 26. Harmonic Distortion at BF in the Cochlea. The total harmonic distortion at BF is at least -30dB for all input rms amplitudes.

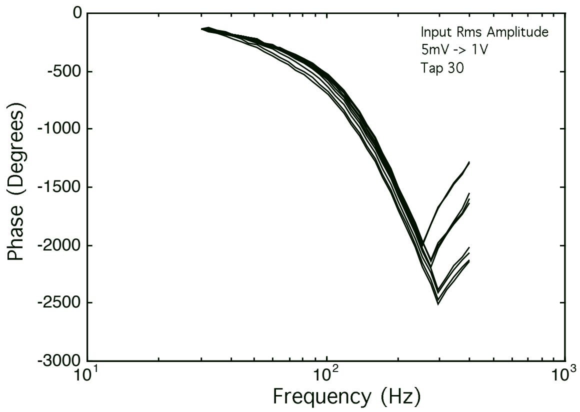

Fig. 27. Phase Characteristics of the Cochlea. Because the AGC corrects Q, but not $1/\tau$, there is relatively little shift in the phase curves with input rms amplitude. The discontinuous parts of the phase curves are due to parasitic effects.

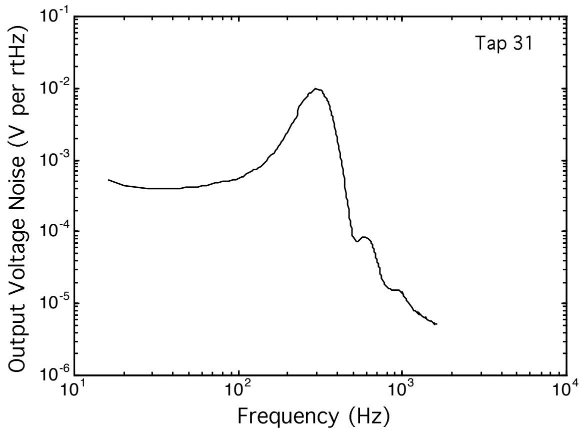

Fig. 28. Typical Noise Spectrum of a Cochlear Tap. The secondary peaks at the high frequencies are at multiples of the primary peak frequency and are due to nonlinearities.

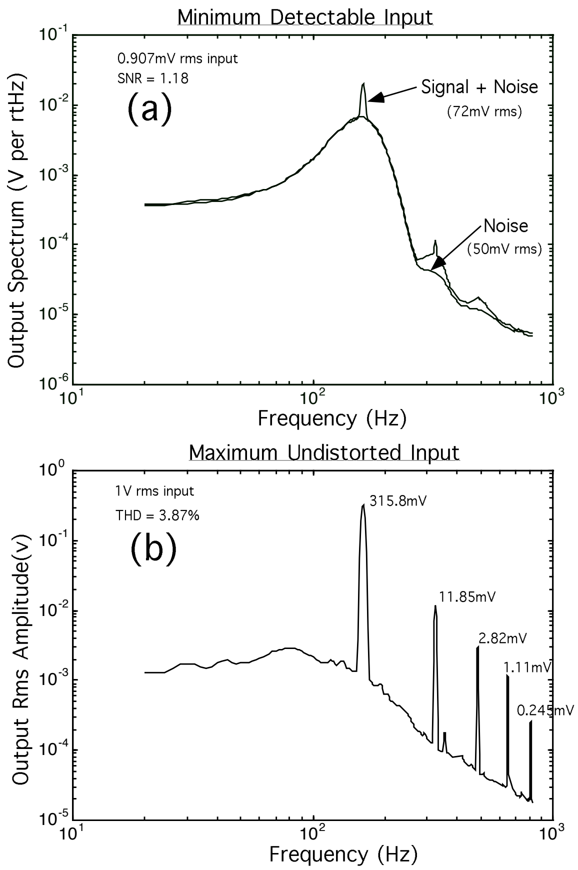

Fig. 29. Dynamic Range of a Cochlear Tap. (a) The spectra of tap 30 when there is no input present, and when a BF signal that is just above the threshold of audibility is present, are shown. The minimum detectable input at BF was found to be 0.875mV. (b) The total harmonic distortion from all harmonics for a 1V rms input at BF was less than 4%. The maximum undistorted input is thus 1V.

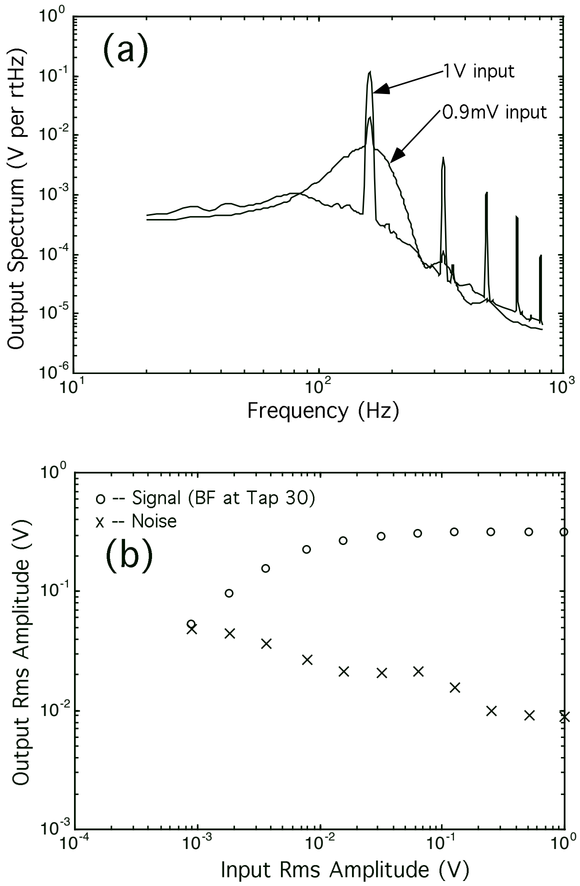

Fig. 30. Signal-and-Noise Amplitude Characteristics. (a) The output spectrum of tap 30 for a 1V rms and 0.9mV rms input at BF shows the adaptation in Q and consequent reduction in noise. (b) The output rms amplitude of the signal and of the noise at different input rms amplitudes are shown.

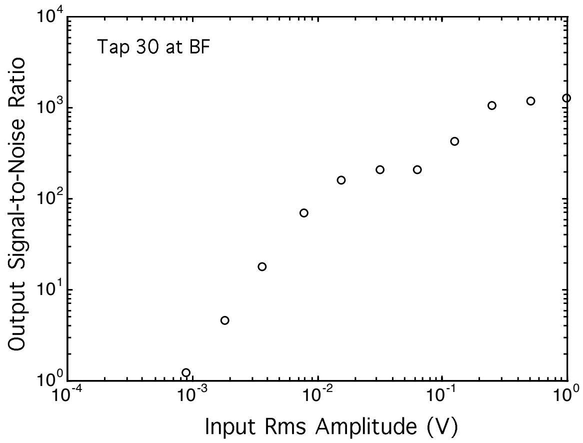

Fig. 31. SNR Amplitude Characteristics. The output SNR improves by about 30dB (1 to 1241) as the signal changes intensity by about 60dB (0.9mV to 1V)

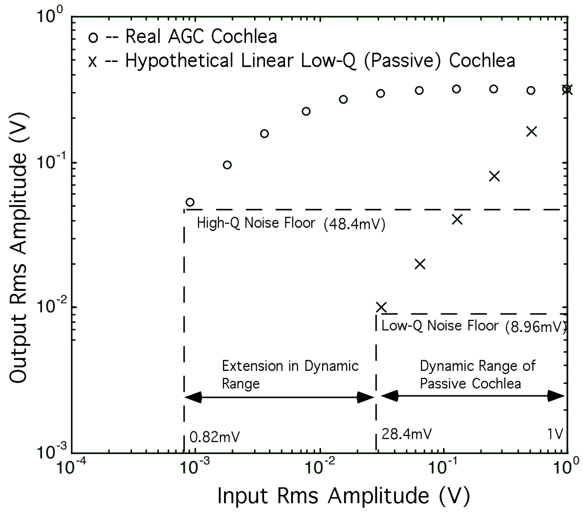

Fig. 32. Extension of Dynamic Range. A hypothetical low-Q cochlea that is completely linear would have a dynamic range of only 30dB due to the uniformly low gain of 0.315 at all amplitudes; such a cochlea is analogous to the passive biological cochlea with no outer hair cells. Our AGC cochlea has a dynamic range of 60dB because faint signals at 0.82mV are amplified by a factor of 59 to be just above the noise floor of 48.4mV, whereas loud signals at 1V rms amplitude are attenuated by a factor of 0.315 to prevent distortion; such a cochlea is analogous to the active cochlea with outer hair cells.

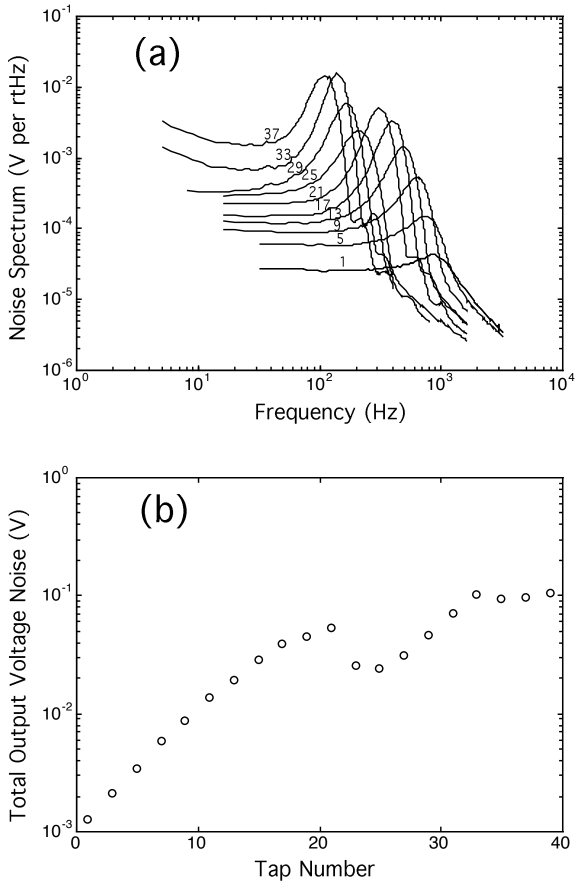

Fig. 33. Noise Accumulation Across Cochlear Taps. (a) The noise spectra at various cochlear taps are shown. (b) The total output noise integrated over all frequencies asymptotically converges due to the exponential taper of the cochlea. The discontinuities in the curve are due to the discontinuous reduction in bandwidth, in turn due to chip offsets, between taps 21 and 25 in (a)

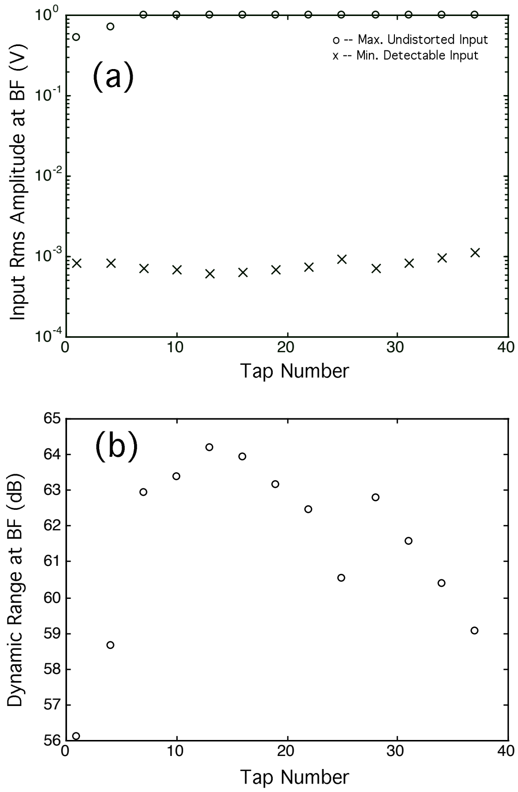

Fig. 34. Dynamic Range Across Cochlear Taps. (a) The minimum detectable input and maximum undistorted input at various cochlear taps are shown. (b) The dynamic range at BF at various cochlear taps are shown.

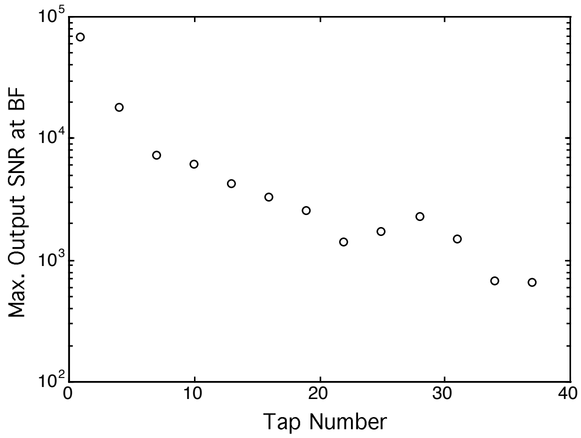

Fig. 35. SNR Across Cochlear Taps. The maximum output signal-to-noise ratio progressively decreases as we travel down the cochlea due to the accumulation of noise. The numbers represent the ratio of the signal power to the noise power, that is 10^5 corresponds to 50dB.

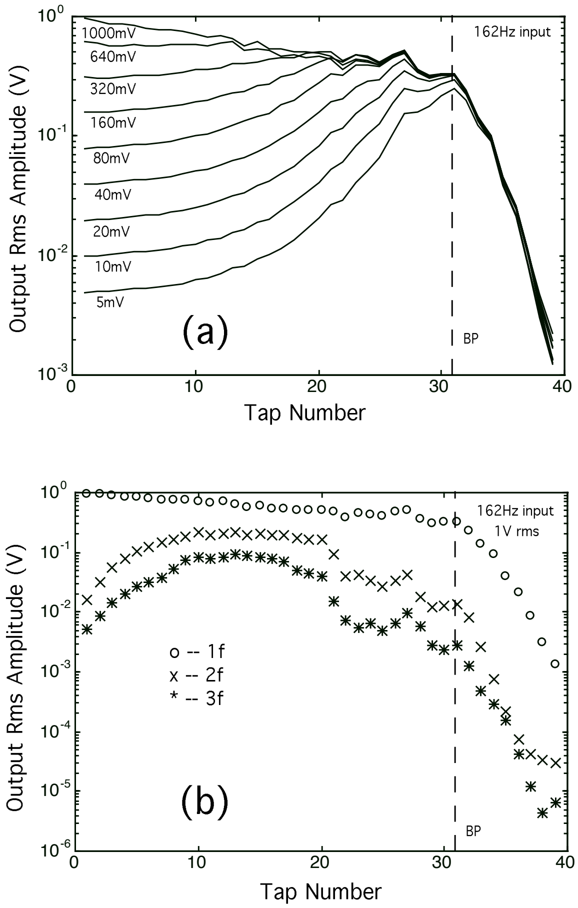

Fig. 36. Spatial-Response Characteristics. (a) The spatial response at various input amplitudes is remarkably similar to the frequency response at various input amplitudes because of the cochlear frequency-to-place transformation ($\log(f) \rightarrow x$). (b) The harmonic distortion is filtered sharply at the best place (BP); it is at its worst at a place that has a corner frequency that is 1 octave above that of the best-place corner frequency.

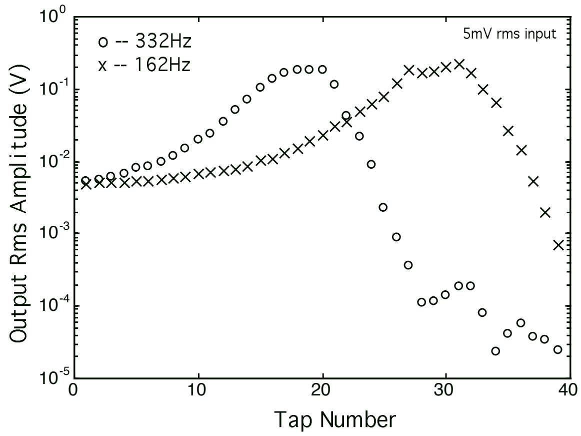

Fig. 37. The Frequency-to-Place Transformation. The best place for high frequencies occurs earlier than that for low frequencies.

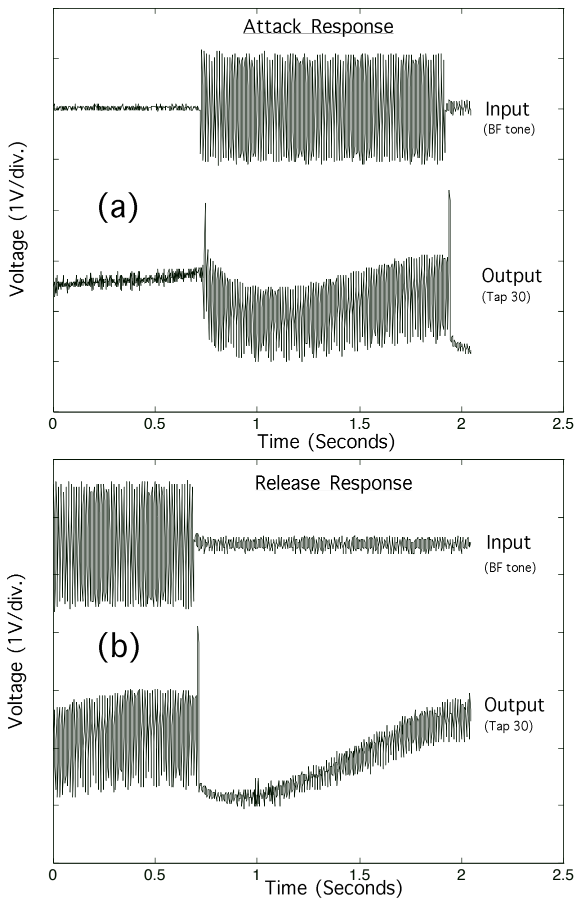

Fig. 38. AGC and Offset Adaptation. (a) At the onset of a loud input tone after a period of silence, there is a brief output transient followed by quick adaptation of the envelope. The offset adaptation occurs in parallel with the envelope adaptation, which happens on a much slower time scale. (b) The reduction in the intensity of a loud input tone causes a gradual adaptation in the envelope of the signal. The offset adaptation is still slower than the envelope adaptation, but the time scales are more comparable.

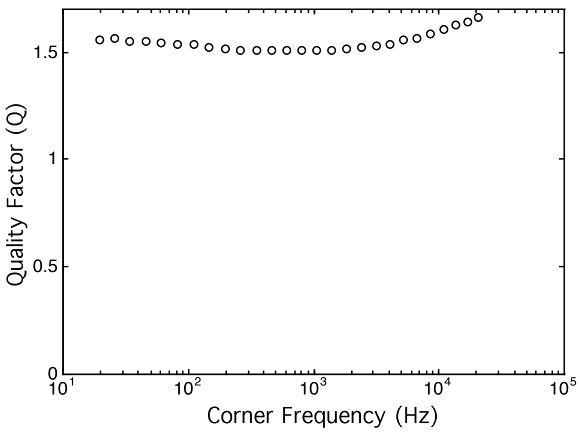

Fig. 39. The Q across Cochlear Filters. The Q across various cochlear taps is fairly well matched.

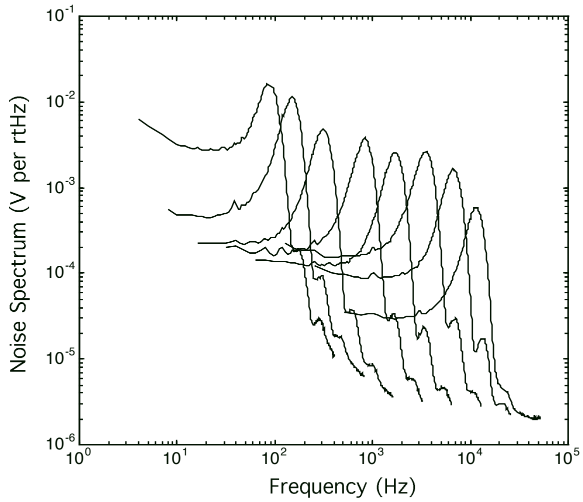

Fig. 40. The Sounds of Silence. The noise spectra at various cochlear taps from the low, mid, and high-frequency cochleas are fairly similar in shape.

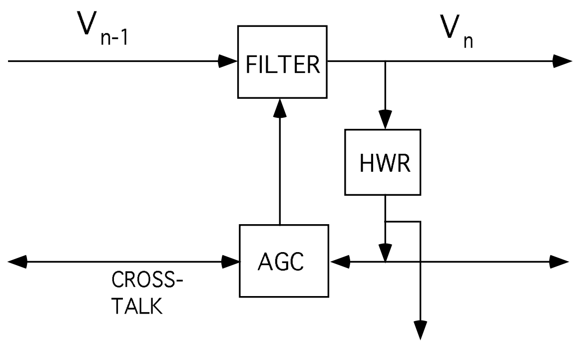

Fig. 41. The ASIC Digital Cochlea

Rahul Sarpeshkar double majored in Electrical Engineering and Physics at MIT between the years of 1986 and 1990. He interned at the T.J. Watson Research Center, IBM as an undergraduate. He obtained his PhD in Computation and Neural Systems at the California Institute of Technology in Carver Mead's analog VLSI lab in 1997. He is currently a member of the technical staff of Bell Laboratories in the department of biological computation. He has done work on noise in devices, circuits, and systems, on adaptive silicon cochleas, and in visual motion detection. His interests lie in analog circuit design for sensory systems, and in the fundamental limits of computing due to physics. His bachelor's thesis won an award at MIT. He is a member of Tau Beta Pi and Sigma Pi Sigma.

Richard F. Lyon received the B. S. degree in engineering and applied science from California Institute of Technology in 1974 and the M. S. degree in electrical engineering from Stanford University in 1975. He has worked on a variety of projects involving communication and information theory, digital system design, analog and digital signal processing, VLSI design and methodologies, and sensory perception at Caltech, Bell Labs, Jet Propulsion Laboratory, Stanford Telecommunications Inc., Xerox PARC, and Schlumberger Palo Alto Research. In his current position as Distinguished Scientist at Apple Computer, he does research in handwriting recognition, neural networks, and models of hearing. Dick is also in his 13th year on the Computation and Neural Systems faculty at Caltech, where he works on sensory modeling research and analog VLSI techniques for signal processing.

Carver A. Mead, Gordon and Betty Moore Professor of Engineering and Applied Science, has taught at the California Institute of Technology for more than 30 years. He has contributed in the fields of solid-state electronics and the management of complexity in the design of very large scale integrated circuits, and has been active in the development of innovative design methodologies for VLSI. He has written, with Lynn Conway, the standard text for VLSI design, *Introduction to VLSI Systems*. His recent work is concerned with modeling neuronal structures, such as the the cochlea and retina using analog VLSI systems. His newest book on this topic, *Analog VLSI and Neural Systems*, was published in 1989 by Addison-Wesley. Professor Mead is a member of the National Academy of Sciences, the National Academy of Engineering, the American Academy of Arts and Sciences, a foreign member of the Royal Swedish Academy of Engineering Sciences, a Fellow of the American Physical Society, a Fellow of the IEEE, and a Life Fellow of the Franklin Institute. He is also the recipient of a number of awards, including the Centennial Medal of the IEEE.