AUTOMATIC GAIN CONTROL IN COCHLEAR MECHANICS

Richard F. Lyon

Advanced Technology Group
Apple Computer, Inc., Cupertino, CA 95014, USA
and Computer Science Department
California Institute of Technology, Pasadena, CA 91125, USA

Introduction

Measurements of basilar membrane motion show that the cochlea has a strong compressive nonlinearity over a wide range of sound intensities, even down to low intensities where the system might be expected to be linear. Many models of cochlear hydrodynamics and micro-mechanics ignore this strong nonlinearity in order to be able to apply linear systems concepts, sometimes resulting in inappropriate interpretations of cochlear function. We propose a modeling approach based on explicitly recognizing the purpose of the strong nonlinearity as an automatic gain control (AGC) that serves to map a huge dynamic range of physical stimuli into the limited dynamic range of nerve firings.

We discuss two aspects of AGC in cochlear function. First, we cover possible AGC mechanisms and mathematical modeling techniques. These involve active outer hair cells whose properties are controlled by the efferent system, resulting in variable-gain wave propagation in the cochlea.

Second, we consider the implications of AGC-based modeling on two controversial issues in the field of cochlear function and modeling: sharpness of neural and mechanical tuning, and two-tone suppression. The single most important implication is that the linearized transfer function of the cochlea at any particular sound intensity is always much broader than an iso-response tuning curve; this fact depends very little on the other modeling details. More detailed features of iso-response tuning curve shapes (such as a notch between the tip and the tail) may depend on details of the AGC system more than on details of the transfer functions.

The physiology of medial olivocochlear efferents as mediators of a wide-dynamic-range spatially coupled binaural AGC system is discussed briefly. We propose using contralateral suppressor tones to directly assess the gain reduction effect of the efferents.

Finally, we summarize our view of cochlear function involving variable negative damping wave mechanics resulting in unsharp pseudoresonant transfer functions.

Background

By AGC we mean a mechanism for varying the sensitivity or gain of a system based on the signal level at the output of the system, so as to reduce the dynamic range of the output relative to the input. A single-input single-output system with AGC is sometimes referred to as a compressor. An AGC generally has a "gain-control loop" with a lowpass "loop filter", so that the system gain varies slowly, rather than instantaneously. An extreme form of AGC attempts to keep the output level constant, as in an AM radio receiver—we do not interpret the term AGC as implying this extreme. An AGC may either vary a "pure gain" or vary some parameter that effects an approximate gain variation coupled with other changes, such as the change of tuning sharpness in the cochlea.

Evidence for a variable-gain mechanism in cochlear function was clear in the pioneering work of Rose et al. (1971). Their data suggest "the existence of a cochlear sensitivity control mechanism which may, but perhaps need not be, mechanical in nature." Subsequent experiments to measure the mechanics of basilar membrane motion

directly bear out the notion that much of the sensitivity variation occurs in the mechanical gain; the work of Robles et al. (1986) shows this effect most clearly, even below 20 dB SPL.

An early model of gain control in cochlear mechanics was proposed by Kim et al. (1973), using an instantaneously-varying damping element. This approach works fairly well for modeling many effects, and is appealing for its simplicity; it results in response characteristics similar to bandpass-nonlinearity (BPNL) models, without requiring a second filter.

In his more recent work, Kim (1984) proposes "biomechanical gain control" mediated by the efferent large medial olivocochlear neurons that innervate the outer hair cells. He provides a detailed review of anatomical and physiological evidence for his hypotheses concerning the roles of the hair-cell and nerve populations. This paper may be taken as the basis of our current views on cochlear mechanics and the need for an AGC that goes beyond the effects of the instantaneous nonlinearity.

Gain-Control Mechanisms and Models

There is plenty of evidence that the relationship between mechanical motion at the oval window and mechanical motion of the basilar membrane is nonlinear over a wide range of signal levels, with apparent gain changes in excess of a factor of 100 (40 dB) for sine waves near CF. Apparently, an active amplification of weak traveling waves gradually changes to an attenuation of strong traveling waves as the input level is increased. The coupling of the mechanical wave propagation system to the active electro-mechano-chemical system of the outer hair cells is generally presumed to provide the mechanism for the variable gain, though details remain to be worked out. Other mechanisms may also provide gain variation, as in the stapedial reflex changing the gain of the middle ear, and depletion of neurotransmitters changing the transduction gain of the inner hair cells. Mechanisms at all levels of the sensory nervous system cooperate to map a wide range of real-world signals into more-or-less "invariant" neural representations, but in this paper we focus only on the gain variations that occur in the mechanical propagation of waves in the cochlea.

Wave propagation in the cochlea can be described by a relationship between complex wavenumber, place, and frequency, when the system is behaving linearly. Nonlinear and adaptive effects can be included in the wavenumber relations for a linearized approximation that applies to a particular state of the parameters (e.g. a particular gain vs. frequency setting that is the result of a particular steady-state stimulus). This type of analysis is easiest if there are no instantaneous nonlinearities, so that the system is short-time linear. The complex wavenumber encodes the spatially varying wavelength as well as the varying energy gain or loss seen by the traveling wave. A varying negative imaginary part of the wavenumber yields a varying-height gain peak in the overall "pseudoresonant" transfer function from the input to any chosen place. Varying the peak gain by changing the wavenumber will inherently vary the shape of the transfer function as well, though the shape and position of the peak may change only slightly compared to the large peak height variations.

Physically-based models of the cochlear mechanics generally involve expressions for the force exerted on the fluid by the cochlear partition (basilar membrane), sometimes in the form of a complex impedance. The force may include terms for membrane stiffness, tension, mass, and losses. A loss term that exerts a force proportional to membrane velocity may model viscous loss in a boundary layer at the membrane, at least approximately. Varying the magnitude of the loss term, and in particular letting it be negative, is typically the way to include a variable gain to model active and adaptive outer hair cells. Terms other than the loss may also be used to provide an active gain, as in Mountain's (1983) "active-stiffness" model.

Since frequencies above CF need to be attenuated, it is important that the loss term change from negative to positive near CF. If a pure negative damping (or negative viscosity) term is used, another loss mechanism must be introduced to attenuate high frequencies. Lyon and Mead (1988) have proposed a hypothetical force proportional to the rate of change of longitudinal curvature of the partition, which provides a loss that rises steeply with frequency or wavenumber and eventually overwhelms the gain from the negative viscosity term. Mountain's active-stiffness model, on the other hand, uses a first-order lowpass function in converting displacement to force, so that for low

frequencies it is a stiffness, and for frequencies above its corner it provides nearly ninety degrees of phase shift to convert the force to an active gain; the magnitude of the active gain declines with frequency, so that a fixed passive loss term eventually dominates at high frequencies. Whether either of these models accurately describes the physics of the cochlear partition with active outer hair cells is not yet clear, but both seem to be at least qualitatively reasonable.

In modeling cochlear hydrodynamics with an active gain mechanism, the nature of the nonlinear relationship between signal level and gain is still a huge open question. Modeling both the static and dynamic aspects of the gain adaptation can be quite challenging, and there is often not enough experimental data to support a choice of model parameters. In many cases, modelers elect to forego the complexity of a dynamic gain-control loop, and settle for an instantaneous compressive nonlinearity in the active gain mechanism. This leaves the efferent system without a functional role in the model, which does not seem reasonable. An understanding of cochlear function will almost certainly require attention to both types of nonlinear compression.

Implications for Sharpness and Suppression

The single most important implication of AGC in the cochlea is the difference in sharpness between (sharp) iso-response tuning curves and (unsharp) iso-intensity response curves or equivalent linearized transfer functions. This difference is clear in much of the experimental data in the field of hearing, whether measured in cochlear mechanics or in nerve responses. However, this difference in sharpness has not generally been interpreted as the simple result of an AGC, and indeed many modelers still make the mistake of trying to match model transfer functions to iso-response data.

Mechanical data of Robles et al. (1986) from the Mössbauer technique clearly indicate the effect of nonlinear amplitude compression on tuning curve sharpness. The BM input-output level functions of their Figure 1 emphasize the amplitude compression for frequencies near CF, rather than tuning. Our Figure 1 shows the same data as two types of tuning curves. Notice that the iso-velocity curve is considerably sharper, by any measure, than even the sharpest low-level transfer function, and much sharper than the higher-level transfer functions that are in effect over much of the range of measurement. We have shown the same effect in simulation of nonresonant models of cochlear mechanics with adaptive negative damping as the AGC (Lyon and Mead 1988).

Iso-response curves have been the most popular way to characterize tuning sharpness for many years. Because iso-response measurements require input levels changing over many orders of magnitude, gain control mechanisms are pushed to their extremes. For frequencies off CF, the input level must be raised to compensate for the falling transfer function gain, and raised more to counteract the gain reduction caused by raising the input level; hence, the skirts are pushed up sharply. Very near the tip, however, the system may in some cases operate below the level that starts to cause a gain reduction. Then within a few dB of the tip, the shapes should agree fairly closely. Indeed, Evans (1977) reported good agreement in the shapes of iso-rate tuning curves and revcor-derived transfer functions, but only within 10 dB of the tips and only for the lowest stimulus levels. Some modelers have accepted his data as evidence that these two types of curves should agree, but this interpretation is especially suspect in the light of clearly-established input-output nonlinearities.

In addition to single-tone data, Robles et al. (1986) present input-output relations in the presence of a second (suppressor) tone above CF. The tone above CF does not cause much response itself, but reduces the gain to tones near CF by more than 10 dB. When the suppressor is significantly more intense than the probe tone, it determines the gain of the system, resulting in a linear input-output relationship to the weaker probe tone in a region where the system was nonlinear when the probe tone was presented alone. These suppression effects are exactly as would be expected from an AGC-based model.

For suppressor tones below CF, two-tone suppression is still very important (mechanical data are not yet published on this, but Ruggero says it's true). Models with instantaneously-varying damping elements do not reproduce this effect, but models with AGC can, under appropriate conditions. In order for this effect to occur, it is necessary that the gain-control feedback from places responsive to the lower frequency also serve to reduce the gain at more basal places, which are sensitive to the higher frequency. We have referred to this cross-place or cross-channel inhibition as "coupled AGC" (Lyon

1982). In the wave propagation model, coupling from a high-frequency place to a low-frequency place is inherent, since waves must travel through the former to the latter. Coupling in the other direction must be provided by having the gain-control feedback signals connect to places more basal than the places whose output they are responsive to. This same cross-coupling toward the base results in significant steepening of the low-side skirt of the iso-response curves, and may contribute the notch below CF that is sometimes modeled as a micro-mechanical antiresonance.

Models with per-channel uncoupled AGC mechanisms will not do at all well in reproducing sharp tuning curves or two-tone suppression; a "lateral" gain reduction is crucial in modeling these effects.

The data in Figure 1 are not as complete as we would like, and our approximated curves are subject to significant error. Better measurements from the laser Doppler velocimeter technique will soon be available to resolve remaining uncertainties. In particular, by using both transient (click) and steady-state (tone and noise) stimuli, it should be possible to separate out the compressive effect due to instantaneous nonlinearities from that due to an efferent-mediated gain-control loop. It should also be possible to measure the small-signal gain reduction in one cochlea due to a suppressor presented contralaterally, presuming there is such an effect.

Active outer hair cells have often been invoked to explain tuning sharpness in excess of that implied by linearized passive models of cochlear mechanics. From the above discussion we hope to make it clear that active gain is not sufficient to produce realistic tuning curves—adaptive compressive control of the gain is required in addition. Both sharp pure-tone tuning and two-tone suppression must be the result of nonlinear compression in cochlear mechanics, whether that compression results from instantaneous nonlinear elements or efferent feedback or both.

When an AGC-based modeling approach is used, sharp enough iso-response curves can be obtained without postulating a resonance based on a large basilar membrane mass; the iso-velocity data of Robles et al. (1986) is reasonably well matched using a two-dimensional wave analysis and zero mass (Lyon and Mead 1988). Contrast this approach to the use of long-wave one-dimensional analysis with a large membrane mass, which has been popular because the transfer functions are sharp enough to match iso-response data directly, ignoring nonlinearities.

Physiology of Cochlear AGC

Liberman (1986) has reported that efferent fibers in the cochlear nerve are typically almost as sharply tuned as afferents, but some are tuned for ipsilateral tones and some are tuned for contralateral tones. In addition, these fibers often have a dynamic range between threshold and saturation of 70 dB or more. He also found that these efferent fibers innervate several outer hair cells near the place corresponding to their CF. All of these observations are exactly what would be expected for the feedback path that controls the mechanical gain of the cochlea, presuming that the outer hair cells provide a controllable negative damping. Let us consider these observations one at a time.

1. The relatively small number of efferent fibers probably carry amplitude information aggregated from groups of the more numerous afferent fibers, thereby causing a slight smearing of the tuning curves.

2. The gains of the two cochleas are probably coupled, so the inter-aural intensity differences are not compressed as much as average intensities, to facilitate using such cues for localization. Such coupling would imply tuning to sounds in either ear, whether via separate units or binaurally-sensitive units.

3. The range over which the firing rates of the efferent fibers should be non-saturating should be about the same as the range of intensities for which the efferent-mediated compressive nonlinearity is active. This scheme is much more economical of fibers than the multiple levels of thresholds of quickly-saturating units often seen in the afferent system.

4. Fibers that inhibit the gain only very near the region of the CF of the corresponding afferents (whose activity is being used to control the gain) would yield a minimum of spatial coupling. A more broadly-coupled AGC is probably preferred, so that local spectral contrasts will not be compressed so much. Coupling the efferent to places more basal than the corresponding afferents will also give more control of the gain, since signals being detected travel through that region. Coupling toward the base also gives

suppression of tones at CF by lower frequencies, and steepens the low-side skirt of tuning curves, perhaps even adding a "notch" below CF. The observed spatial spread of up to nearly an octave in distance seems reasonable.

Measurements of cochlear mechanical response in the presence of a contralateral suppressor tone would be a good way to directly assess efferent-mediated gain control.

A View of Cochlear Function

Building on the AGC-based modeling approach, we have assembled a view of cochlear function that we believe is reasonably self-consistent and in agreement with much of the experimental data in the field. Many of the pieces of this view are not generally accepted, and indeed do not appear to be correct within modeling paradigms without AGC and too heavily dependent of linear analysis. In particular, the following items characterize our currents views:

1. Near CF, the cochlea operates mainly in the short-wave region, rather than in the long-wave region. The long-wave region does not contribute significantly to neural response, except perhaps in the low-frequency tails.

2. The mass of the cochlear partition is negligible, except perhaps near the base, and the tension of the cochlear partition is also negligible. Membrane stiffness and fluid mass interact to form a dispersive wave propagation system without resonance.

3. The hydrodynamic system of the cochlea is not highly tuned (in the sense of being highly frequency selective or resonant), and the best frequency for a place is quite level dependent.

4. Sharp iso-response tuning curves are the result of an AGC operating in conjunction with a broadly tuned hydrodynamic system—no "second filter" or other tuned sharpening mechanism is needed to model cochlear tuning as seen in either the mechanics or the neural firings.

5. A bandpass filterbank to model cochlear response ought to be designed as a cascade, rather than as a parallel bank of independent filters, in order to achieve a realistic amplitude and group-delay response, to model propagation of distortion products, and to conserve on computation.

6. A model of the active adaptive cochlea must be extremely nonlinear over a wide range of signal level, mainly to effect dynamic range compression.

7. Odd-order distortion products in the hydrodynamic system are audible under special conditions, and are maximum for cochlear input power levels comparable to or a little higher than the power that can be supplied by the outer hair cells.

8. Viscosity of the cochlear fluid is negligible except at a boundary layer at the basilar membrane. That is, very little energy is dissipated in the bulk of the fluid, but as the wave collapses onto the membrane in the short-wave region, energy is dissipated into the membrane and into fluid drag along the membrane.

9. Outer hair cells act approximately like a negative boundary layer viscosity, adding a force to the membrane proportional to its velocity, for frequencies below CF. The transition from negative damping to positive damping is not sharply tuned.

10. In the normally-functioning cochlea, energy travels in one direction—standing waves, acoustic emissions, and reflections may be neglected except in pathologies.

This approach has been the basis of our work on cochlear modeling applied to speech recognition, including custom VLSI digital and analog implementations of cochlear models.

Conclusions

The progress of scientific research is generally accelerated through the use of good models to guide the design and interpretation of experiments. A key element missing from most auditory models in recent years is a dynamic AGC. We believe it is now important to include such an AGC element to help resolve some of the controversies in the hearing field, such as questions of sharpness of tuning. We propose that neurophysiologists might do well to devote more effort to answering the many open questions concerning the details of cochlear mechanical gain variation, in order to contribute to the development of a comprehensive and realistic model of hearing.

References

Evans, E.F. (1977), "Frequency Selectivity at High Signal Levels of Single Units in Cochlear Nerve and Nucleus." In: *Psychophysics and Physiology of Hearing* (Eds: Evans, E.F. and Wilson, J.P.) Academic Press, London, pp. 185–192.

Kim, D.O., Molnar, C.E. and Pfeiffer, R.R. (1973), "A system of nonlinear differential equations modeling basilar-membrane motion." J. Acoust. Soc. Am. **54:**1517–1529.

Kim, D.O. (1984). "Functional Roles of the Inner- and Outer-Hair Cell Subsystems in the Cochlea and Brainstem." In: *Hearing Science* (Ed: Berlin, C.) College-Hill Press, San Diego, pp. 241–261.

Liberman, M.C., and Brown, M.C. (1986), "Physiology and anatomy of single olivocochlear neurons in the cat." *Hearing Res.* **24**:17–36.

Lyon, R.F. (1982), "A Computational Model of Filtering, Detection, and Compression in the Cochlea." *Intl. Conf. on Acoust., Speech, and Sig. Proc.*, IEEE, Paris, pp. 1282–1285.

Lyon, R.F. and Mead, C. (1988), "An Analog Electronic Cochlea." *IEEE Trans. Acoust., Speech, and Sig. Proc.* **36**: 1119–1134.

Mountain, D.C., Hubbard, A.E., and McMullen, T.A. (1983), "Electromechanical Processes in the Cochlea," In: *Mechanics of Hearing* (Eds: de Boer, E. and Viergever, M.A.) Martinus Nijhoff, The Hague, pp. 119–126.

Robles L., Ruggero, M.A. and Rich, N.C. (1986). "Mössbauer measurements of the mechanical response to single-tone and two-tone stimuli at the base of the chinchilla cochlea." In: *Peripheral Auditory Mechanisms* (Eds: Allen, J.B. et al.) Springer-Verlag, New York, pp. 121–128.
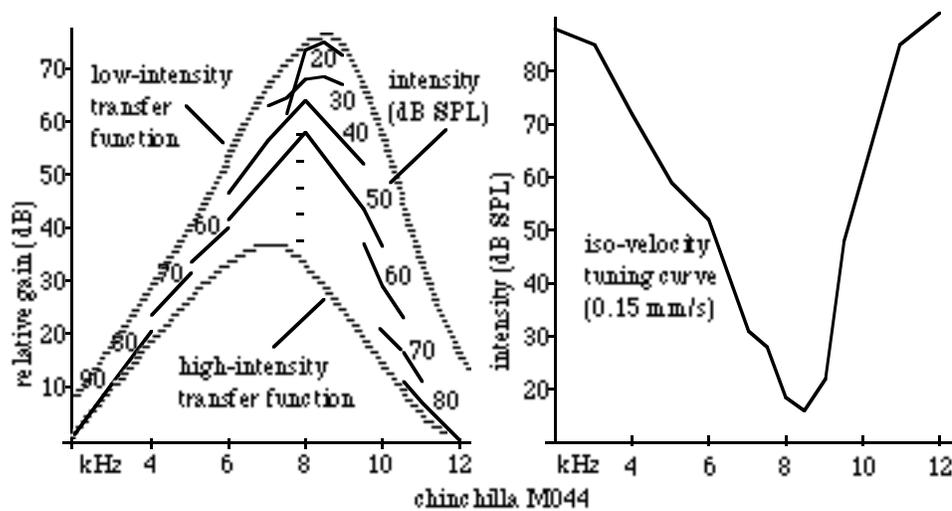
FIGURE 1   Iso-intensity data and approximated curves (left) and iso-velocity data (right) re-plotted from Figure 1 of Robles *et al*. (1986). Notice that the iso-intensity data show an input-output compression of about 2:1 near CF; that is, each 10 dB increase in input intensity causes a 5 dB decrease in the mechanical gain of the cochlea. This compression results in the iso-velocity curve being significantly sharper than even the sharpest of the transfer functions.