The EarSpring Model for the Loudness Response in Unimpaired Human Hearing

David McClain, Refined Audiometrics Laboratory, LLC

December 2006

Abstract

We describe a simple nonlinear differential equation for a harmonic oscillator with power dependent stiffness that adequately describes the main features of unimpaired human loudness perception. Solutions of the model are presented, and a mechanism for relating model vibrational power to Sones loudness measure is described. This model is used for estimating hearing correction targets from impaired levels to what would be heard by unimpaired hearing.

Contents

1 Introduction **1**

2 Qualitative Description of the Model **2**

3 Equation of the Model **4**

4 Estimation of Model Parameters **5**

5 Frequency Dependence **9**

6 A Model for HyperRecruitment **9**

1 Introduction

EarSpring is the name for our model for the whole-system human loudness response for unimpaired hearing. It is an ad-hoc empirically driven model that captures the essential features of human hearing. It is not specifically a cochlear model. Rather, it models the

entire system of hearing, incorporating the combined effects of chochlea, afferent 8th nerve, brain, and efferent 8th nerve components, and their complex interactions.

Our aim is to understand how unimpaired human hearing responds to signals at various intensity levels. We utilize this normal hearing loudness response to help derive needed corrections for impaired listeners. EarSpring shows us how loudly sounds ought to appear. And so the model furnishes correction targets.

We utilize empirical measurements to establish values for the model parameters based on:

- The Sones ratio between the hearing threshold level and the reference level of 40 dBSPL at 1 kHz.
- The detuning (flattening) of tones by 75 cents as tones range in intensity from 40 dBSPL to 90 dBSPL near 1 kHz.

EarSpring is able to successfully demonstrate the main features of loudness perception:

- Linear response to near-threshold level sounds, although for normal daily environmental sound levels this is of little importance.
- Nearly cube-root behavior for more common sound levels, with intensities above 40 dBSPL near 1 kHz.
- Computer simulations of the EarSpring equation demonstrate the production of odd harmonic distortion products from loud signals, in agreement with Inter-Modulation Distortion (IMD) products measured with loud bass-frequency excitation signals.

2 Qualitative Description of the Model

To within experimental accuracy, and to first order, we find that the loudness perception at any one pitch, of unimpaired human hearing, can be adequately described by the measure of average vibration power in a nonlinear damped harmonic oscillator.

The nonlinearity is provided by a spring constant whose stiffness grows with increasing vibration power. We propose that its nominal resonant frequency, the pitch sensation it produces, is that frequency at which the oscillator resonates at very low power levels, where its behavior is essentially the same as a simple linear damped harmonic oscillator.

As input sound levels grow in intensity, the stiffness of the model's spring constant increases in proportion to the vibration power of the induced oscillations. At high enough input levels this nonlinearity begins to dominate the loudness response.

The result is essentially a linear damped harmonic oscillator for threshold level sounds, and which produces a cube-root response over sound levels in the range of common daily sound intensities above 40 dBSPL.

As sound intensity increases, and the stiffness constant grows, the natural resonant frequency of the oscillator shifts toward higher frequencies, and moves above the excitation frequency. This shifting produces two effects: loudness compression and pitch shifting.

As the natural resonant frequency of the oscillator shifts above the excitation frequency, the degree of absorbed excitation diminishes with increasing sound intensity. This gives rise to the compression observed as cube-root, wherein for a doubling of apparent sound loudness we require an increase of 10 dB in sound intensity.

And if we conceptualize the pitch perception of human hearing as corresponding to a parallel bank of such oscillators, those in adjacent frequency bands gradually shift their own resonant frequencies higher in response to increasing sound intensity. And so the lower bands, whose stimulation corresponds to some lower pitch in our sensation, become increasingly excited at higher sound intensities. As a result one can imagine that the excitation will produce an apparent pitch that is increasingly flatter than what the excitation would produce at low intensity levels.

Both of these effects are readily measured in the laboratory. The cube-root behavior is well known and readily verified for common sound levels above 40 dBSPL. The pitch flattening has been measured as approximately 75 cents for an overall sound intensity increase from 40 dBSPL to 90 dBSPL, near 1 kHz.

Threshold region sounds, those below 40 dBSPL, are uncommon in normal daily experience, and can generally only be heard well in isolation booths. While important for demonstrating the linearity of human loudness response at very low sound intensity levels, this response is not very important for the description of common daily hearing experiences.

We further postulate that our model is the same for all frequency bands when expressed in terms of Phon excitation. The ear canal modifies the sensitivity of our hearing as sounds range in frequency due to resonant enhancement for sounds near 4 kHz. Higher order modifications occur as a result of pinnea and head shadowing directional effects. Transmission of sound into the cochlea and its coupling to the basilar membrane also varies in efficiency with frequency.

Our assumption is that there is a basic mechanism at work, whose description at different frequencies produces similar results for similar excitation levels reaching the EarSpring. Hence, in absolute terms, and to first order, one must first apply an Equal-Loudness correction to absolute intensities expressed in dBSPL, obtain equivalent excitation levels as expressed in Phon, before applying the results of our model.

Bear in mind that this is a whole-system model, not a cochlear model. It is not known specifically where such a spring constant could be located, how the oscillator's power is sensed, nor what might produce the rising stiffness with increasing excitation levels. It is conceivable that a good portion of this behavior does reside in the cochlea, where power perception might be aided by stimulation of inner hair cells in response to vibration of the basilar membrane, and where some of the increasing stiffness may arise from afferent nerve stimulation of outer hair cells to diminish the flexibility of the basilar membrane.

But we are not so much interested in cochlea specific mechanisms, but rather in the overall loudness response. Cochlear behavior is only one component of the entire system composed of cochlea - efferent nerves - brain - afferent nerves. Hearing is the result of individual behaviors from, and interactions between, the subcomponents of this system. We find that our model is the simplest that adequately describes this whole system.

3 Equation of the Model

The EarSpring model is a second order nonlinear differential equation relating the amplitude of vibration of a mechanical system to a driving force. We view the hearing mechanism as a parallel bank of EarSpring systems, each tuned to a different resonant frequency.

The equation is very nearly the same as that for a conventional damped harmonic oscillator, except that it incorporates a stiffness term which increases in magnitude with the average power of vibration:

$$\left[\frac{d^2}{dt^2} + 2\beta\frac{d}{dt} + k\left(1 + \gamma\left\langle y^2\right\rangle\right)\right] y(t) = F(t)$$

where:

t	=	time
$y(t)$	=	amplitude of vibration
$F(t)$	=	driving force
$\left\langle y^2\right\rangle$	=	mean power of vibration
β	=	damping constant
k	=	spring constant
γ	=	coefficient of power dependence of spring constant

For any fixed level of excitation, the average power of vibration $\left\langle y^2\right\rangle$ in the steady state will be a constant. Hence, we see that this describes a harmonic oscillator with poles in its transfer function at $i\beta \pm \omega_0$, where $\omega_0 = \sqrt{k_\gamma - \beta^2}$, and $k_\gamma = k\left(1 + \gamma\left\langle y^2\right\rangle\right)$:

$$H(\omega) = \frac{1}{-\omega^2 + 2i\beta\omega + k\left(1 + \gamma\left\langle y^2\right\rangle\right)}$$

While increasing damping β moves the poles off the real frequency axis and decreases the natural resonant frequency, an increase in the power of vibration moves the poles, in a direction parallel to the frequency axis, toward frequencies further from the origin. Spring stiffness increases, which raises the resonant frequency.

4 Estimation of Model Parameters

In the following discussion we will view our system as being excited by a sinusoidal force at the undamped natural frequency $\sqrt{k}$. We can make a frequency independent equation by dividing the EarSpring equation through by k. We will be measuring the system response at this same frequency, and so the reduced transfer function, at that frequency, becomes:

$$H(\omega_0 = \sqrt{k}) = \frac{1}{2i\hat{\beta} + \gamma\left\langle y^2\right\rangle}$$

where, now, $\hat{\beta}$ is the frequency normalized damping constant ($\hat{\beta} = \beta/\sqrt{k}$).

Since increasing vibrational amplitude shifts the resonant frequency of the EarSpring system, excitation applied with increasing intensity at the undamped resonant frequency will have a gradually lessened effect as the system resonance moves further from the excitation frequency.

We can estimate the γ parameter on the basis of tone flattening as excitation intensity rises from 40 to 90 dBSPL[1]. We might view tone flattening as the consequence of adjacent channel EarSprings shifting their resonant frequencies higher, thereby causing increasing excitation of the next lower channel, and hence the excitation appears ever flatter.

[1]For frequencies near 1 kHz. Units of dBSPL are inappropriate for use at other frequencies.

We need the excitation presented to the oscillator. And so units of Phon, derived from dBSPL intensities adjusted by a frequency dependent Equal-Loudness correction, are appropriate at all frequencies.

At 1 kHz these systems of measurement are defined such that 40 Phon = 40 dBSPL, a difference of 10 dBSPL equals a difference of 10 Phon, and where 0 dBSPL is the nominal threshold for hearing which corresponds to an RMS pressure amplitude of 0.0002 dyne/cm^2.

At any other frequency the threshold in dBSPL will differ, but will always be denoted as 0 Phon, and the scaling of Phon differences to dBSPL differences will also be different.

In frequency normalized terms, the expression for the resonant frequency of our EarSpring transfer function is:

$$\hat{\omega} = \sqrt{(1 + \gamma \langle y^2 \rangle) - \hat{\beta}^2}$$

Taking the ratio of this equation for the two different excitation levels and then solving for the γ power term gives us:

$$\Gamma_{40} \equiv \frac{1}{2}\gamma |a_{40}|^2 = \frac{\left(F_{90}^2 - 1\right)}{\left(S_{90} - F_{90}^2\right)} \left(1 - \hat{\beta}^2\right)$$

where,

$$\begin{aligned} a_{40}, \hat{\omega}_{40} &= \text{vibration amplitude and resonant frequency at 40 dBSPL driving force} \\ a_{90}, \hat{\omega}_{90} &= \text{vibration amplitude and resonant frequency at 90 dBSPL driving force} \\ S_{90} &\equiv |a_{90}|^2 / |a_{40}|^2 \\ F_{90} &\equiv \hat{\omega}_{90} / \hat{\omega}_{40} \end{aligned}$$

and where we have assumed a sinusoidal driving force. The leading $\frac{1}{2}$ factor arises from the fact that the mean power of a sinusoid of amplitude a is proportional to $\frac{1}{2}|a|^2$.

The pitch flattening ratio F_{90} has been measured in the lab as approximately 75 cents ≈ 1.044.

We don't yet know the individual values for a_{40} and a_{90}, but we can estimate their magnitude ratio by using the approximate cube-root behavior of loudness sensation with rising sound intensity. Hence a difference of 50 dB in sound intensities should produce a ratio $S_{90} \approx 17\,\text{dB}$.

From these expressions, we can utilize the ratio of the transfer function at two input levels to find the amplitude of vibration for any other input signal level in terms of its ratio to a_{40} as:

$$S_P \equiv \frac{|a_P|^2}{|a_{40}|^2} = \left(\frac{P}{100 P_{thr}}\right)^2 \frac{\left(4\hat{\beta}^2 + \Gamma_{40}^2\right)}{\left(4\hat{\beta}^2 + (\Gamma_{40} S_P)^2\right)}$$

where P is the excitation pressure level, and P_{thr} is the threshold pressure level corresponding to 0 dBSPL. This is a cubic equation in the amplitude magnitude-squared ratio, S_P.

Benade[2] reports that the ratio S_0 for threshold level stimuli, where $P = P_{thr}$, ought to be $1/(22)^2 \approx 0.002$. So, from the two equations above, we can solve for $\hat{\beta}$. With this initial estimate for $\hat{\beta}$ we can solve for an improved ratio S_{90}, and then iterate this process by recomputing Γ_{40} and then $\hat{\beta}$ until we converge to the value for $\hat{\beta}$.

Once the value of $\hat{\beta}$ has been found, we can map the amplitude magnitude ratio S_P over all input dBSPL levels and obtain the solution shown in Figure 1.

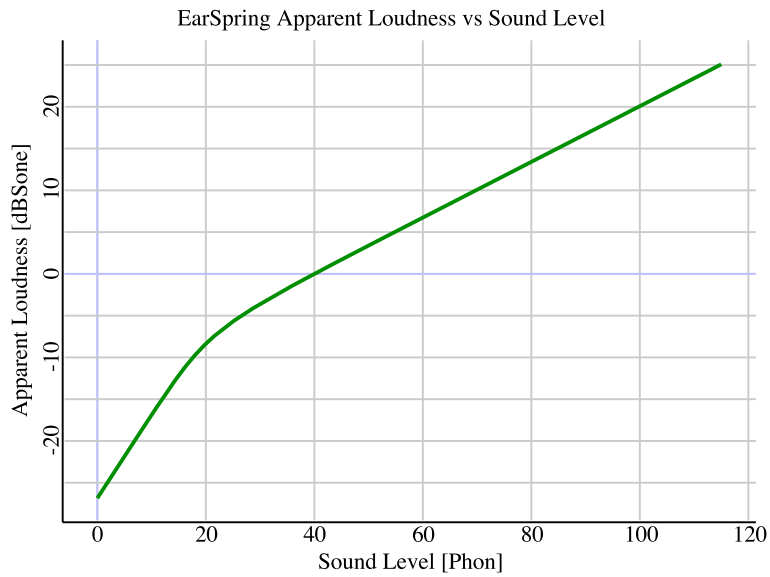

Figure 1: *EarSpring Amplitude versus Driving Force. Linear behavior is apparent near threshold level intensities. At normal sound levels above 40 dBSPL we see cube-root compression.*

As we see, the behavior of the system near threshold sound levels is nearly linear. At sound intensities above 40 dBSPL, where most normal daily sounds are found, the behavior is nearly cube-root. The average slope of the graphed solution, over the range from 40 to 100 dBSPL, is $\frac{1}{2}\alpha \approx 0.335$.[3]

[2]Benade, Arthur H., *Fundamentals of Musical Acoustics - Second, Revised Edition*, Dover, 1990, pg. 252.

[3]See the next page for the meaning of α.

Figure 2 shows the slope of the curve in Figure 1. Near threshold levels its slope is very close to unity, indicating linear behavior. At sound intensity levels much above 40 dBSPL, the slope has settled to a nearly constant value of 1/3.

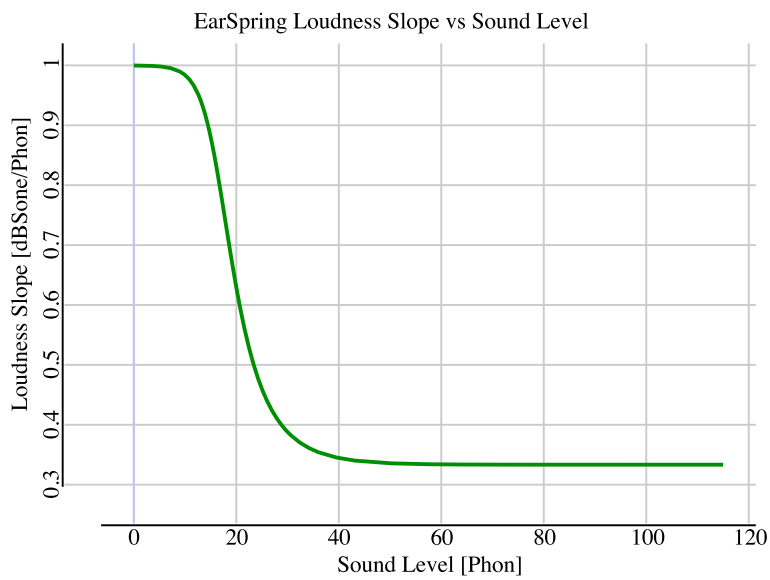

Figure 2: *EarSpring Slope versus Driving Force. These slope values are* 1/2 *of the characteristic exponent* α *discussed in the text.*

We make the identification of the EarSpring vibration power ratio, S_P, with the Sones level for each source intensity level. At threshold sound levels Benade[4] gives the expression for Sones as:

$$\text{Sones} = \left(\frac{P}{22P_{thr}}\right)^2 \qquad \text{for } P \approx P_{thr}\text{, near threshold levels}$$

Exponent α arises in the expression for Sones at intensity levels above 40 dBSPL:

$$\text{Sones} = \left(\frac{P}{100P_{thr}}\right)^\alpha \qquad \text{for } P/P_{thr} > 40 \text{ dBSPL}$$

Benade gives the value for $\alpha \approx 0.6$, but this value is an average over many individuals whose measured values ranged anywhere from 0.5 to 0.7. From our EarSpring model we derive an average value for $\alpha \approx 0.669$ over the intensity range from 40 to 100 dBSPL.

[4] *ibid*

5 Frequency Dependence

The identification of Sones with the EarSpring vibrational power ratio S_P versus the sound intensity P in dBSPL can only be correct at 1 kHz. The Sones scale is calibrated as 1 Sone for a 40 dBSPL sinewave at 1 kHz.

At other frequencies, the outer ear modifies the base relationship in a manner first shown by Fletcher and Munson. Unimpaired hearing is more sensitive near 4 kHz than anywhere else, and this arises from the resonant enhancement provided by the ear canal.

Hence to be more correct, the driving force in the EarSpring equation should be identified with units of Phons. These units, defined as identical to dBSPL at 1 kHz, vary in both their zero point and scaling relative to dBSPL at other frequencies. But iso-Phon contours in the dBSPL-Frequency plane correspond to the sensation of equal loudness, and so we make the assumption that these units also describe the degree of coupling of sound to our EarSpring oscillators. When EarSpring is viewed in terms of excitation levels presented in this manner the solution obtained above becomes independent of frequency.

In order to utilize the EarSpring solution for hearing correction targets, we must first estimate the sound intensity level in each frequency band in dBSPL. We then apply the frequency dependent Equal-Loudness correction to these sound intensities to derive Phon levels above threshold presented to the EarSpring. After this conversion the EarSpring model can furnish sensible target apparent loudness levels, in Sones, for use in hearing corrections.

6 A Model for HyperRecruitment

It should be clear that if the γ term in the model stiffness were too small, the suppression of excessive vibration amplitude at high sound field intensities becomes impaired. Figure 3 shows the consequence of a γ term that is too small by a factor of 20.

The essential feature of HyperRecruitment at high sound-field intensities is the perception of sound being too loud. That is shown by the height of the red curve above the nominal unimpaired hearing curve shown in green.

Such an insufficiency in γ might be caused by the deterioration of outer hair cells in the cochlea, or a failure of the brain to control vibrational amplitudes of the basilar membrane through excitation of the outer hair cells.

Diminished hearing, through the destruction of inner hair cells alone, is not able to account for HyperRecruitment because at high enough sound-field intensities the sensation

of sound produced by even a limited number of inner hair cells should be sufficient to correct excessive basilar membrane vibration amplitudes.

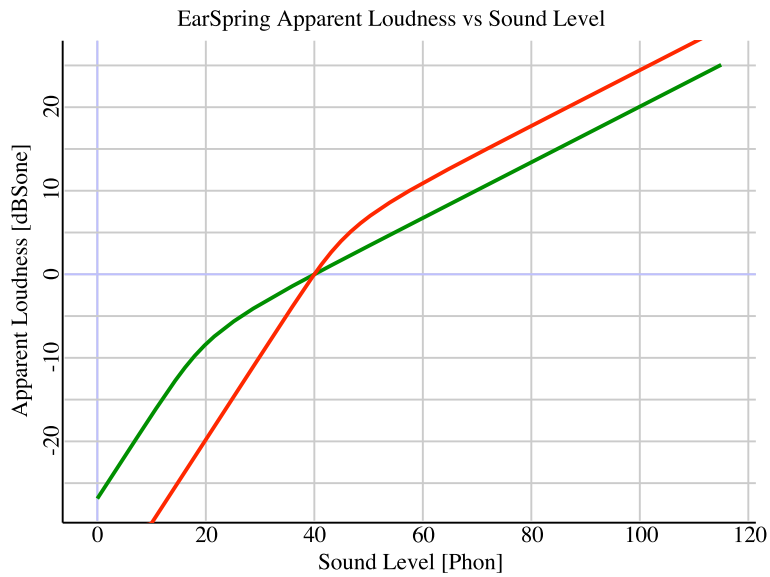

Figure 3: *EarSpring model for HyperRecruitment induced by an insufficiency in the γ term of the model stiffness. Unimpaired hearing response in green, HyperRecruitment in red.*

We suggest that a γ insufficiency is only part of the mechanism for HyperRecruitment. Another paper discusses the more general cases and models for Recruitment, HyperRecruitment, and Decruitment, and their correction to EarSpring target behaviors.