isvr Institute of Sound and Vibration Research

Models for Compressive Nonlinearities in the Cochlea

S. J. Elliott and J.M. Harte

ISVR Technical Memorandum 913

May 2003

University of Southampton

SCIENTIFIC PUBLICATIONS BY THE ISVR

Technical Reports are published to promote timely dissemination of research results by ISVR personnel. This medium permits more detailed presentation than is usually acceptable for scientific journals. Responsibility for both the content and any opinions expressed rests entirely with the author(s).

Technical Memoranda are produced to enable the early or preliminary release of information by ISVR personnel where such release is deemed to the appropriate. Information contained in these memoranda may be incomplete, or form part of a continuing programme; this should be borne in mind when using or quoting from these documents.

Contract Reports are produced to record the results of scientific work carried out for sponsors, under contract. The ISVR treats these reports as confidential to sponsors and does not make them available for general circulation. Individual sponsors may, however, authorize subsequent release of the material.

UNIVERSITY OF SOUTHAMPTON

INSTITUTE OF SOUND AND VIBRATION RESEARCH

SIGNAL PROCESSING & CONTROL GROUP

Models for Compressive Nonlinearities in the Cochlea

by

S J Elliott and J M Harte

ISVR Technical Memorandum No. 913

May 2003

Authorised for issue by
Prof S J Elliott
Group Chairman

CONTENTS

FIGURES

Page

Abstract

The response of the basilar membrane in the cochlea displays a compressive nonlinearity, which is conventionally described using an input-output level curve for sinusoidal excitation. Several models of this nonlinear behaviour are considered, in particular an instantaneous power-law nonlinearity, and two systems with level-dependent properties, one with an automatic gain control and the other a dynamic system having amplitude-dependent damping. It is shown that the parameters of each of these models can be chosen so that they give a level curve with a slope of 1 dB/dB at low levels and a slope of ½ dB/dB at higher levels, as observed in the response of the basilar membrane.

In order to try to distinguish between the responses of these rather different models their distortion and transient responses are also considered, but these are found to be surprisingly similar. The instantaneous input-output characteristics of these models are clearly different, however. The instantaneous nonlinearity has a single-valued but nonlinear characteristic, whereas the level-dependent systems have an almost linear characteristic, for a given amplitude of excitation, whose slope varies with the excitation level.

1. Introduction

There are several sources of nonlinearity within the ear. This report concentrates on a nonlinearity within the cochlea associated with the compressive nonlinearity in the response of the basilar membrane, which is caused by saturation of the active processes which sharpens its frequency response at low levels (Pickles, 1988). There are several ways of describing and characterising this nonlinearity, but one widely-used representation is a graph of the log amplitude of the basilar membrane (BM) motion against the log amplitude of the sinusoidal driving pressure: the input-output level curve. If the excitation frequency is close to the frequency of maximum BM response at low levels (the characteristic frequency or best frequency), then the BM level curves directly measured in a healthy cochlea by a number of authors (Sellick et al., 1982; Nuttall and Dolan, 1996; Rhode and Recio, 2000) show some similar characteristics. At sound pressure levels below about 30 dB the level of the BM response rises with sound pressure level at a slope of about 1 dB/dB, indicating that the BM is responding linearly at these low levels. Above a sound pressure level of about 30 dB, the slope of the level curve decreases, typically to about ½ dB/dB, which indicates a compressive nonlinearity. This change in slope occurs over a range of about 10 dB to 20 dB in level. At very high pressure levels there is some evidence for another increase in the slope of the level curve, but in this report we will be less concerned with the cochlear response at these high sound pressure levels, than in the transition from linear to nonlinear response at about 30 dB.

As well as causing compression in the BM response, the active processes in the cochlea are also believed to give rise to otoacoustic emissions of various types. The levels of these otoacoustic emissions do not generally increase in direct proportion to the excitation level. Norton and Neely (1987) present data for tonebursts which indicate that the level of the otoacoustic emission rises at about 0.6 dB/dB compared with the excitation level. The derived level curves for distortion product otoacoustic emissions (Dorn et al., 2000) also show a slope of approximately ½ dB/dB at excitation levels about 30 dB, with evidence at some frequencies for a slope closer to 1 dB/dB at lower levels, as seen in the BM response. A similar linear region for otoacoustic emissions at low levels was observed by Zwicker and Schloth (1985).

The purpose of this report is to discuss various simple models for this cochlear nonlinearity. In particular the characteristics and behaviour of an instantaneously-acting nonlinear function will be contrasted with those of two systems whose properties are level-dependent.

Our emphasis is on models of the response at one position on the BM to sinusoidal excitation at different amplitudes but only at the characteristic frequency. We have been less concerned about the response of these models at different frequencies, since a realistic frequency response can only come from a consideration of the coupled motion of a whole region along the BM (de Boer, 1991).

Instantaneous nonlinearities are discussed in Section 2 and have been used to model cochlear response since the 1960's (Engebretson, 1968; Pfeiffer, 1970), when analogue hardware models were used, with diodes to provide nonlinearities. Digital implementations of instantaneous compressive nonlinearities have been used more recently, for example by Goldstein (1995), and by Meddis et al. (2001) in their dual resonance model. It is shown that systems with suitably selected instantaneous nonlinearities can exhibit a level curve that has a slope of 1 dB/dB up to a specified input level and a slope of ½ dB/dB above this level.

Models of the cochlear with an explicit level-dependent gain (i.e. an automatic gain control or AGC) were investigated by Lyon (1990), and analogue implementations of AGC circuits were used in the silicon hardware implementation of the electronic cochlea of Lyon and Mead (1988). In the quasilinear model of Kanis and de Boer (1993) the value of an impedance varies with the level of the pressure, and so this too constitutes a level-dependent model. Section 3 considers the steady-state and transient response of such a level-dependent system. It is shown that if the gain depends in a specific way on the output level, the AGC system also displays a level curve that has a slope of 1 dB/dB up to a specified input level and a slope of ½ dB/dB above this level.

In Section 4 dynamic, mass-spring-damper, models of the BM response are considered, in which the damping is a function of the amplitude of response. The gain of such a system is implicitly level-dependent. These second-order models can be described by Van der Pol's equation, which displays a wide variety of dynamic behaviour, including

limit cycle oscillations, whose properties have striking similarities to those of spontaneous otoacoustic emissions. Such models also display a Hopf bifurcation, from a stable equilibrium to a limit cycle oscillation of increasing amplitude, as the damping parameter is varied. There has recently been interest in using theoretical models of this type to describe the cochlear function (Equíluz et al., 2000; Camalet et al., 2000). One of the predictions of this model is that when the system is just stable, the amplitude of the forced response of the system is proportional to the amplitude of the forcing to the power ⅓, which would result in a ⅓ dB/dB level characteristic. It is shown in Section 4 how these equations can be modified to yield a 1 dB/dB level characteristic at low levels and a ½ dB/dB slope at higher levels, consistent with that discussed above.

A variety of models can therefore be used to predict realistic input-output level curves. One method of distinguishing between the responses of these models could be in terms of their transient response, although it is shown that level-dependent AGC circuits can respond almost as quickly as an instantaneous nonlinearity. Both systems can also exhibit similar levels of harmonic distortion. The clearest difference between the responses of these models of BM response may thus be to plot their *instantaneous* input-output characteristics at a variety of input levels. For an instantaneous nonlinearity all these characteristics should fall on the same line and trace out a clearly nonlinear curve. For a level-dependent system, the instantaneous input-output characteristic will be an almost straight line for a given input, but whose slope will depend on the input level. The level-dependent systems thus have an instantaneous output which is a multiple-valued function of their instantaneous input and so their behaviour *cannot* be properly described by a Volterra or Wiener series.

2. Instantaneous Nonlinearities

2.1 *Power-law Nonlinearity*

Systems which display instantaneous nonlinearities are those for which the current output signal, $y(t)$, is a function only of the current input signal, $x(t)$, i.e.

$$y(t) = f\big(x(t)\big). \tag{2.1}$$

A simple power-law nonlinearity (Smoorenburg, 1972) of the form

$$y(t) = x^p(t)\ \text{sign}\big[x(t)\big], \tag{2.2}$$

will provide a compressive effect provided the power, p, is less than unity. Consider the response of such a system to the sinusoidal excitation

$$x(t) = \sqrt{2}\ X\ \cos(\omega_o t), \tag{2.3}$$

where X is the rms input amplitude and ω_o the angular excitation frequency. The output will be a periodic signal with a Fourier series of the form

$$y(t) = \sqrt{2}\sum_{n=1}^{\infty} Y_n\ \cos(n\omega_o t), \tag{2.4}$$

where Y_n is the rms amplitude of the n-th harmonic. Substituting (2.3) into (2.2), allows the output in this case to also be written as

$$y(t) = \big(\sqrt{2}\ X\big)^p\ \cos^p(\omega_o t)\ \text{sign}\big[\cos(\omega_o t)\big], \tag{2.5}$$

where only the term $\big(\sqrt{2}\ X\big)^p$ depends on the input amplitude. The remaining terms on the right hand side are periodic and so have a Fourier series which can be written

$$\cos^p(\omega_o t)\,\mathrm{sign}[\cos(\omega_o t)] = \sum_{n=1}^{\infty} a_n \cos(n\,\omega_o t), \tag{2.6}$$

where a_n is the n-th harmonic component, which is independent of the input amplitude.

Comparing equations (2.4), (2.5) and (2.6), we see that the amplitude of the nth harmonic of the output is equal to

$$Y_n = a_n\, 2^{\frac{p-1}{2}}\, X^p. \tag{2.7}$$

In particular the fundamental component of the output is

$$Y_1 = a_1\, 2^{\frac{p-1}{2}}\, X^p. \tag{2.8}$$

If the level curve for this nonlinearity is obtained, by plotting $20\log_{10} Y_1$ against $20\log_{10} X$, the slope will be p dB/dB. The ratio of the fundamental component of the output to its sinusoidal input amplitude is called the "describing function", which is widely used in control engineering to analyse the behaviour of feedback systems which contain nonlinear elements. Although the response of such a nonlinear element to a sinusoidal excitation will, in general, not be sinusoidal, it is assumed that only the fundamental component of the output is significant in determining the stability and overall performance of the system. The describing function is, in general, the complex ratio of the fundamental component of the nonlinear element's output to its input, and will depend on the level of the input sinusoid in a way that can be readily calculated for various idealised forms of nonlinearity (Ogata, 1970). Conventional frequency-domain methods of analysing the behaviour of the complete feedback system can then be used, except that the value of the describing function, and hence the behaviour of the system, will vary with level.

Figure 2.1 shows the output waveforms of equation (2.1) for sinusoidal excitation when $p = 0$, ½ and 1. When $p = 0$, the output signal is a square wave, but becomes more rounded when $p = ½$ and is sinusoidal for $p = 1$. Figure 2.2 shows the variation with p

of the slope of the level curve, the fundamental amplitude of equation (2.6), a_1, and the total harmonic distortion, in %, which is defined to be

$$\mathrm{THD}(\%) = \left(\frac{\sum_{n=2}^{\infty} |a_n|^2}{\sum_{n=1}^{\infty} |a_n|^2} \right)^{\frac{1}{2}} \times 100\% . \tag{2.9}$$

The total harmonic distortion is about 44% when $p = 0$, corresponding to the distortion of a square wave, falling to about 17% when $p = ½$, and to zero when $p = 1$, when the output is undistorted.

2.2 *Input Scheduled Power-law Nonlinearity*

In order to obtain a level curve which has a dual slope, as discussed in the introduction, it is necessary for the power law to change with the instantaneous value input signal, so that it may be written as $p(x)$, and the output signal is now

$$y(t) = [\alpha x(t)]^{p(x)} \operatorname{sign}[x(t)], \tag{2.10}$$

where α is a constant. As an example of this kind of nonlinearity we will assume that the variation of the power with the instantaneous value of the input signal takes the form

$$p(x) = 0.75 - 0.25 \tanh\left(|x| - \delta\right) \tag{2.11}$$

where δ is a threshold value for x, well below which $p(x) \approx 1$ and well above which $p(x) \approx ½$. The variation of $p(x)$ with the input, x, normalised by δ is plotted in the lower curve of Figure 2.3, in which the value of δ has been taken to be $10^{1.5}$ to ensure that the level curve changed slope at appropriate input level (30 dB). The upper curve shows the instantaneous input-output characteristic of this nonlinearity, where the constant α in equation (2.10) has been set equal to $1/\delta$ to ensure a smooth curve, together with that of a square-root nonlinearity. The harmonic distortion of this nonlinearity is zero at

low amplitudes and rises to about 17% at high amplitudes, as predicted for a power-law nonlinearity with $p = 0.5$.

Figure 2.4 shows the resulting level curve (fundamental output level against input level for a sinusoidal excitation) which, as expected, has a 1 dB/dB slope at low levels and a ½ dB/dB slope at higher levels.

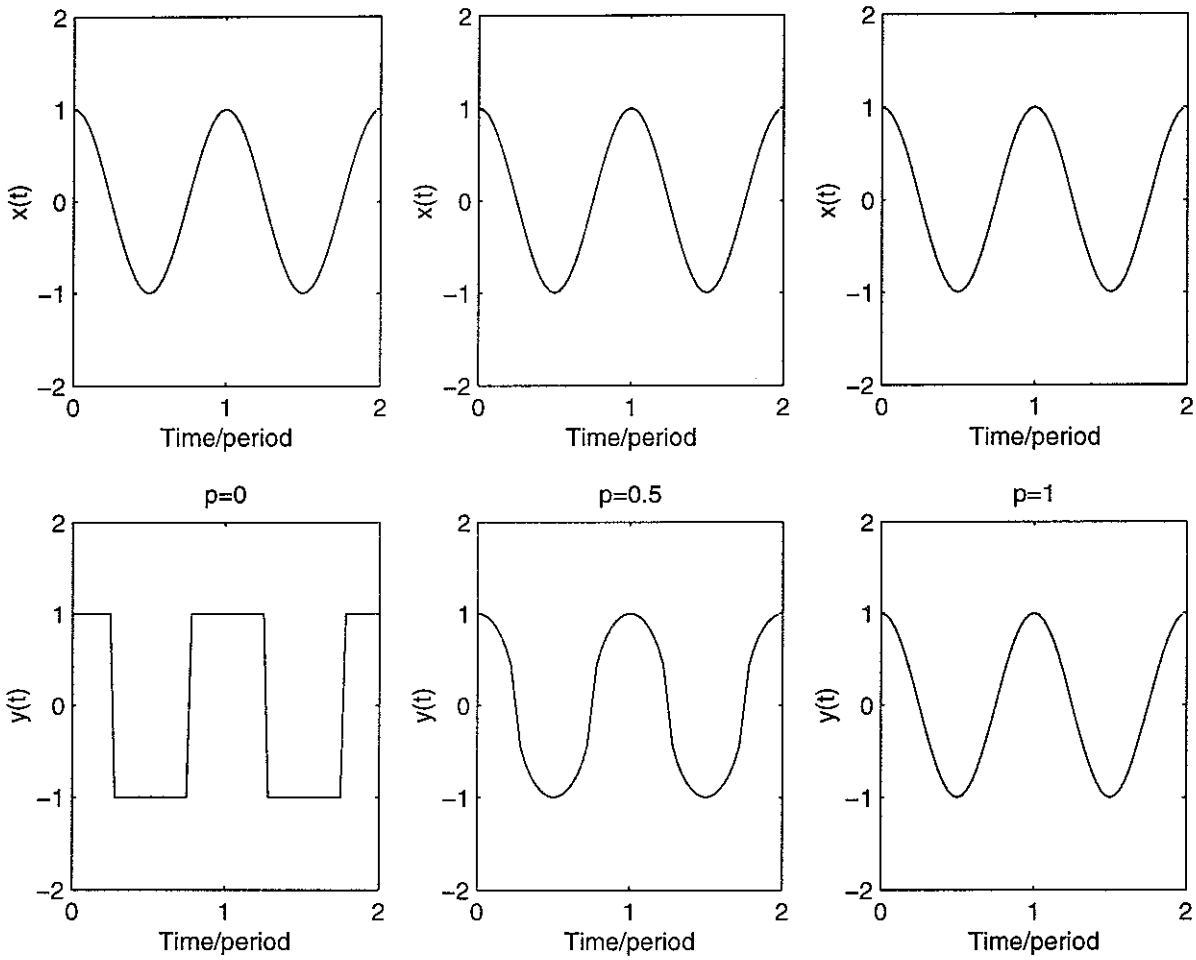

Figure 2.1: The sinusoidal input waveform, $x(t)$, to the power-law nonlinearity (upper) and the resulting output waveform, $y(t)$, (lower), when the power law, p, is 0, ½ and 1.

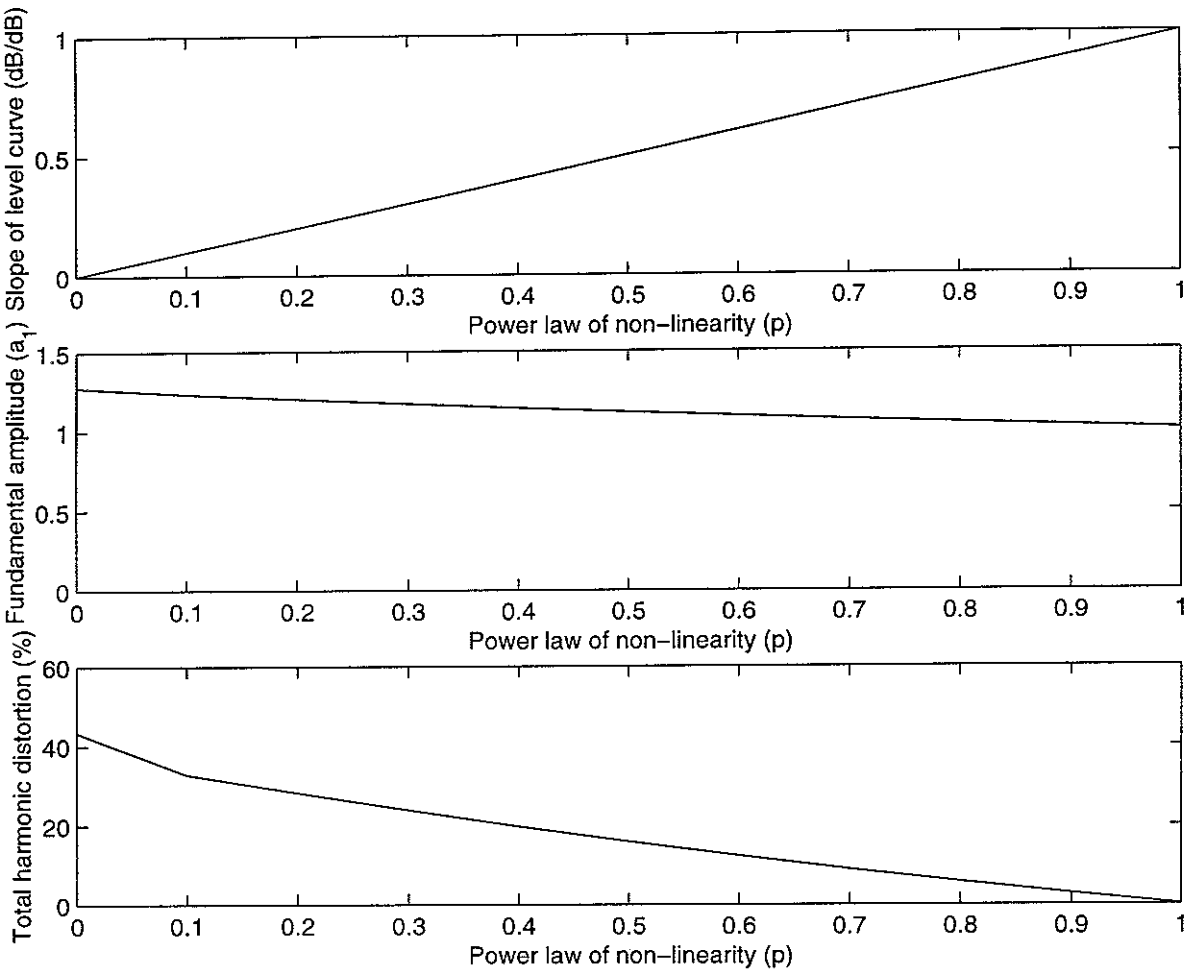

Figure 2.2: The variation with the power-law p of the slope of the level curve, the fundamental amplitude and the percentage total harmonic distortion for sinusoidal excitation of the power-law nonlinearity.

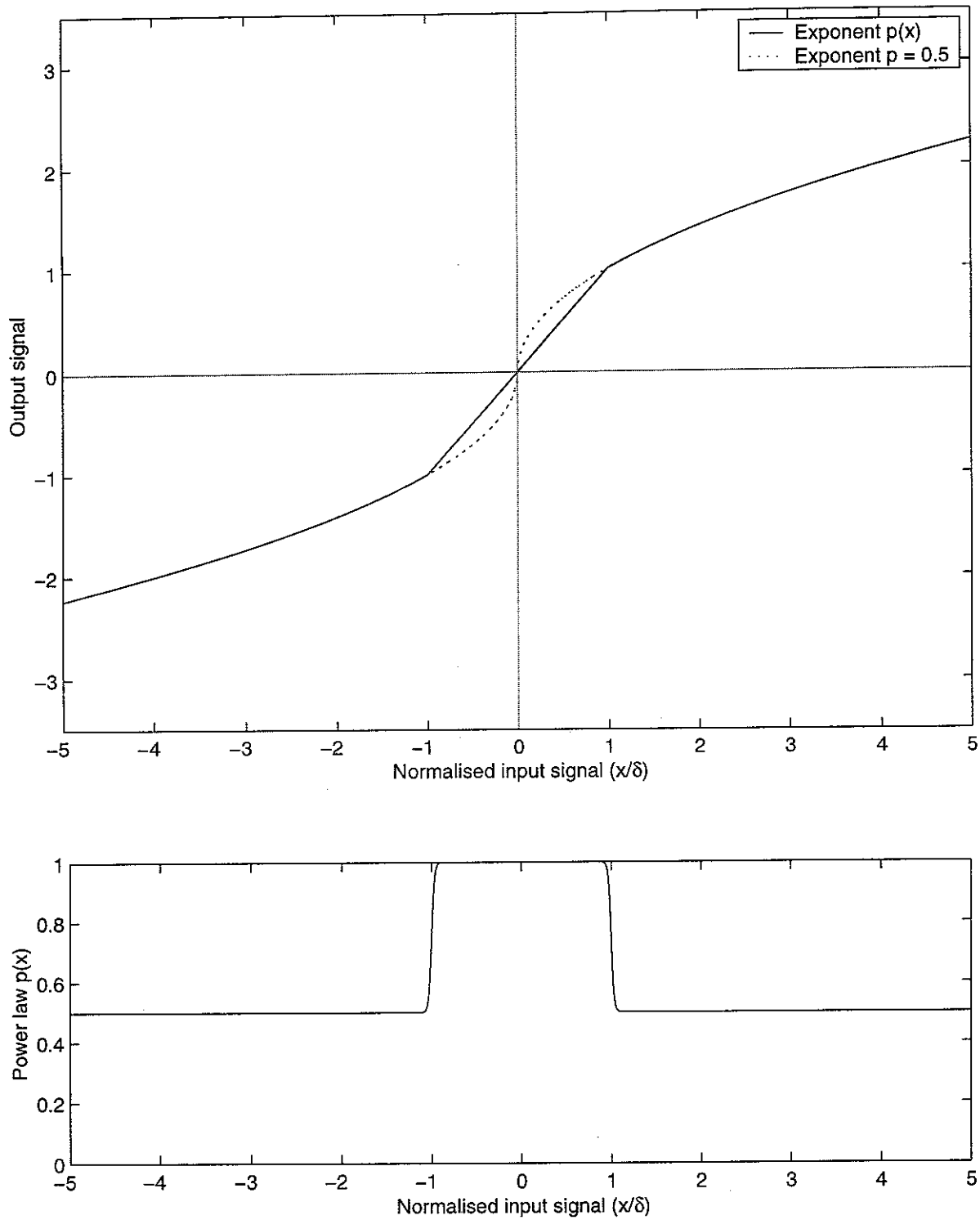

Figure 2.3: The instantaneous input-output characteristic of a nonlinear function (upper curve) in which the power law is scheduled on the instantaneous value of the input signal, as shown in the lower curve.

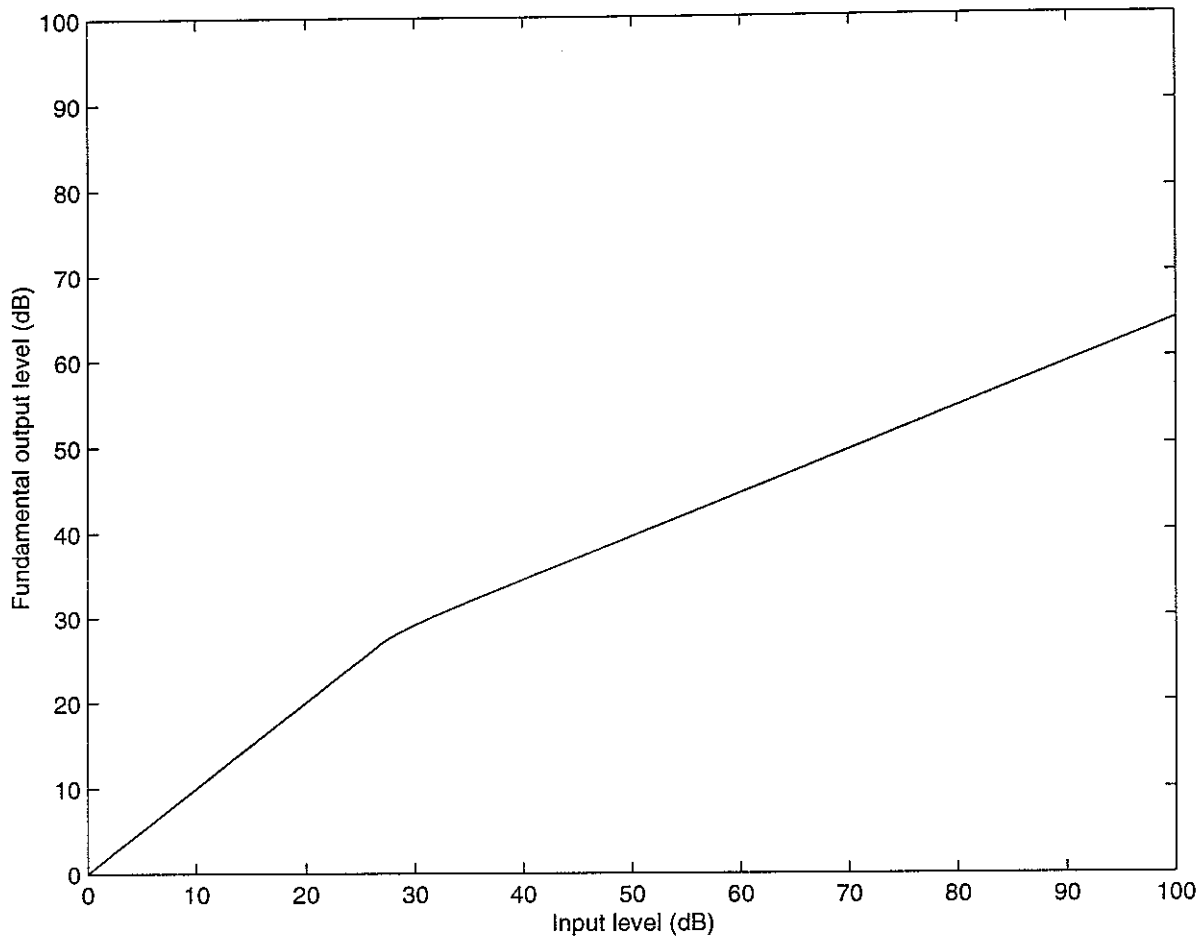

Figure 2.4: The level curve of the instantaneous nonlinearity shown in Figure 2.3.

3. Systems with Level-Dependent Gain

3.1 *The Automatic Gain Control*

One of the simplest forms of level-dependent system which displays nonlinearity is often called an automatic gain control (AGC), and takes the form of a gain, whose value varies with the level of either the input or the output of the system. In this section we investigate the input-output characteristics of a system in which the gain depends on the level of the output, measured using an output detector, as shown in Figure 3.1, excited by a steady-state tonal input. An output-dependent gain has been considered to be a more plausible model of cochlear nonlinearity than an input-dependent gain.

We will assume for now that the detector in Figure 3.1 perfectly measures the rms level of the output signal, and postpone a detailed discussion of its operation until later. If we assume that the sinusoidal input signal is of the form above:

$$x(t) = \sqrt{2}\, X \cos(\omega_o t), \tag{3.1}$$

and that the gain g does not vary during the period of the signal, so that the output signal is undistorted and of the form

$$y(t) = \sqrt{2}\, Y \cos(\omega_o t), \tag{3.2}$$

then the rms output amplitude, Y, is given by

$$Y = g\, X\,. \tag{3.3}$$

In general the gain law may be written as

$$g = \text{function}\,(Y), \tag{3.4}$$

where Y is the rms amplitude of the output. We will initially consider the specific case in which the gain is dependent on the output level raised to a constant power.

$$g = a\,Y^{p}, \tag{3.5}$$

where a is a constant and p is the power. Thus

$$Y = g\;X = a\,Y^{p}X, \tag{3.6}$$

so that

$$Y = (a\,X)^{\frac{1}{1-p}}. \tag{3.7}$$

If, for example, $p = -1$, then $1/(1-p) = \frac{1}{2}$, the level curve of the AGC would have a slope of ½ dB/dB and the system would act as a dynamic range compressor. If $p = +½$, however, then $1/(1-P) = 2$, the slope of the level curve would be 2 dB/dB and the system would act as a dynamic range expander.

Figure 3.2 shows the instantaneous input-output characteristics of a compressive AGC, with $p = -1$, for three different levels of input. Each characteristic is almost a straight line, but the slope changes with the input amplitude, indicating that the gain reduces with level. Since the output signal clearly does not have a single instantaneous value for a given instantaneous value of input signal, such level-dependent nonlinear systems cannot be adequately described by a Volterra or Wiener series. A simple NARMA model of an automatic gain control is given in Appendix A.

3.2 *AGC Level Curves*

If the gain law of an output-level dependent AGC takes the form

$$g = \frac{1}{Y}, \tag{3.8}$$

i.e. the power, p, in the general expression of equation (3.5) is –1, then the gain will tend to infinity as the output level tends to zero, which is physically unrealistic. To avoid this singularity, the gain law can be modified to be of the form

$$g = \frac{1}{Y + \varepsilon}, \tag{3.9}$$

where ε is a constant. Intuitively, if $Y >> \varepsilon$, then the system will behave as in equation (3.8) to give a compressive nonlinearity with a level curve of slope ½ dB/dB. If $Y << \varepsilon$, however, then the gain of the system is independent of input level and the slope of the input-output level curve will be 1 dB/dB.

In quantitative terms, the relationship between the rms amplitudes of the input, X, and output, Y, is given by substituting (3.9) into (3.3) to give

$$Y = \frac{X}{Y + \varepsilon}, \tag{3.10}$$

which may be multiplied out to give a quadratic equation for Y and X. The physically-realistic solution to this quadratic equation takes the form

$$Y = \sqrt{X + \frac{\varepsilon^2}{4}} - \frac{\varepsilon}{2}, \tag{3.11}$$

so that $Y \approx \sqrt{X}$ when $X » \varepsilon^2$, $Y \approx X/\varepsilon$ when $X « \varepsilon^2$ and $Y \approx 0.62\varepsilon$ when $X = \varepsilon^2$. The level curve corresponding to equation (3.11) with $\varepsilon = 10^3$ and an addition gain of 10^3 is plotted in Figure 3.3 and, as expected, has a slope of 1 dB/dB at low levels and of ½ dB/dB at higher levels.

3.3 *Detector Dynamics*

We now consider in more detail the operation of the output detector in Figure 3.1. This is assumed to consist of an instantaneous squaring device, a low pass filter, with frequency response $H(j\omega)$, and an instantaneous square-root device, as shown in Figure 3.4.

The differential equation describing the low pass filter is assumed to be first order and to be of the form

$$u(t) = v(t) + T_L \frac{dv(t)}{dt}, \tag{3.12}$$

where $u(t)$ and $v(t)$ are the input and output signals of the low pass filter, and T_L is its time constant. The frequency response of this system is given by

$$H(j\omega) = \frac{\omega_L}{\omega_L + j\omega}, \tag{3.13}$$

where $\omega_L = 1/T_L$ is the angular cut-off frequency of the filter, so that its cut-off frequency in hertz, f_L, is equal to

$$f_L = \frac{1}{2\pi T_L}. \tag{3.14}$$

If all the alternating components of the low pass filter's input, $u(t) = y^2(t)$, which is the squared AGC output signal, are well above f_L, then the filter will smooth out this signal, to give a constant, d.c., value, which is then used to schedule the gain. In particular if the output is a sinusoid, as above, with an angular frequency, ω_0, which is well above ω_L, then the output of the detector, $d(t)$, will be the rms output amplitude, as expected:

$$d(t) = Y. \tag{3.15}$$

If, however, the input frequency is not much larger than the cut-off frequency of the low pass filter, then the output of the detector, $d(t)$, will have an alternating as well as a steady component. This will cause the gain of the AGC to vary during the course of a cycle of the input signal, thus distorting it and generating harmonics in the output signal. Figure 3.5 shows the waveforms of the various signals in the AGC of Figure 3.1, with a gain law given by equation (3.9), when the input is a sinusoid whose frequency is equal to the cut-off frequency of the AGC's low pass filter. The output

waveform, $y(t)$, is distorted, but has a strong fundamental component, so that its squared waveform, $y^2(t)$, has a dominant component at twice the input frequency. This component is reduced but not eliminated by the low pass filter, so that the detected signal, $d(t)$, is still modulated at twice the input frequency. The low pass filter effectively attenuates any higher harmonic components of $y^2(t)$, due to harmonic distortion in the output signal, $y(t)$. The variation in the detected signal is then reflected in the variation of the gain with time, $g(t)$. Finally, the multiplication of the input signal with the time-varying gain gives rise to the third-harmonic distortion seen in the output waveform.

The phase shift in the harmonic distortion also gives rise to loops in the instantaneous input-output characteristics, which can just be discerned in Figure 3.2. These loops get larger as the input frequency approaches the low pass filter cut-off frequency. As the loops get larger, the instantaneous input output characteristic gets even further from a single-valued nonlinearity.

It is shown in Appendix B that the total harmonic distortion is a function of the ratio of the input frequency, ω_o, to the cut-off frequency of the low pass filter, ω_L. If the normalised input frequency ω_o/ω_L is denoted Ω, then provided Ω is not too small, the total harmonic distortion is approximately equal to

$$\mathrm{THD}(\%) = \frac{1}{8\Omega} \times 100\%, \tag{3.16}$$

which is plotted in Figure 3.6, together with the results of a computer simulation of the AGC with various input frequencies. The theoretical prediction for the total harmonic distortion predicts the simulation results well for normalised excitation frequencies, Ω, above about 2, but cannot predict the more complicated behaviour seen in the simulations for lower values of Ω because the assumptions made in the derivation are no longer valid in this case.

The dynamic behaviour of the low pass filter in the AGC detector also gives rise to an overall transient response of the AGC, which is illustrated in Figure 3.7. This figure

shows the result of a sudden increase in the amplitude of the input signal, $x(t)$, by 20%, and a sudden return to its original value. The upper curve is the input amplitude and the middle curve shows the variation in the gain of the compressive AGC with time. When the input amplitude increases, the gain falls approximately exponentially, with a time constant determined by the low pass filter in the detector, until it reaches a steady state value determined by the gain law. When the input level falls, the gain exponentially increases to its original value with approximately the same time constant in this case. Since the gain cannot change instantaneously with the input amplitude, however, then at the instant that the input amplitude rises by 20%, the gain is at its original value and so the output amplitude also instantaneously increases by 20%. This decays to an increase of about 8% as the AGC gain falls to its steady state value. Similarly when the input amplitude falls, the output amplitude instantaneously falls below its steady state value and then exponentially recovers.

It is shown in Appendix C that for small but abrupt charges in input level, the variation of AGC gain with time is exponential, with a time constant, τ, which is a half that of the low pass filter T_L, so that

$$\tau = \frac{1}{2\omega_L}. \tag{3.17}$$

If the sinusoidal input signal has an angular excitation frequency of ω_o, then its period is equal to

$$T_o = \frac{2\pi}{\omega_o}. \tag{3.18}$$

The time constant of the AGC, in periods of the excitation frequency, is thus predicted to be

$$\frac{\tau}{T_o} = \frac{\omega_o}{4\pi\,\omega_l} = \frac{1}{4\pi}\,\Omega \tag{3.19}$$

where Ω is again the ratio of excitation frequency to the cut-off frequency of the low pass filter. The measured time constants from a computer simulation of the AGC are compared with this theoretical prediction in Figure 3.8. It is difficult to accurately measure the time constant of a decaying sinusoid when the time constant is small compared with the period, but the simulations are in reasonable agreement with the theoretical predictions.

For a given excitation frequency, the choice of the cut-off frequency of the AGC's low pass filter is thus a compromise between distortion, Figure 3.6, and response time, Figure 3.8. Selecting a cut-off frequency of one half the excitation frequency, so that $\Omega = 2$, appears to provide a reasonable compromise between these competing effects, since it gives rise to a distortion of only about 6%, but a time constant which is only of the order of one sixth of the period of the input waveform.

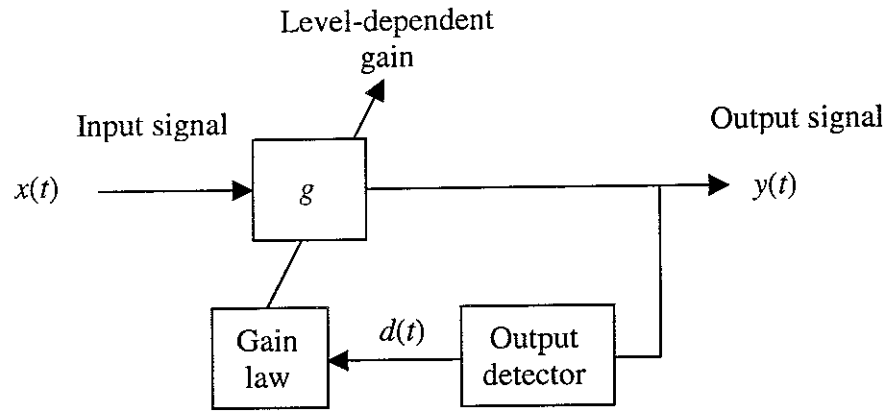

Figure 3.1: Block diagram of a simple automatic gain control whose gain depends on the detected level of the output signal.

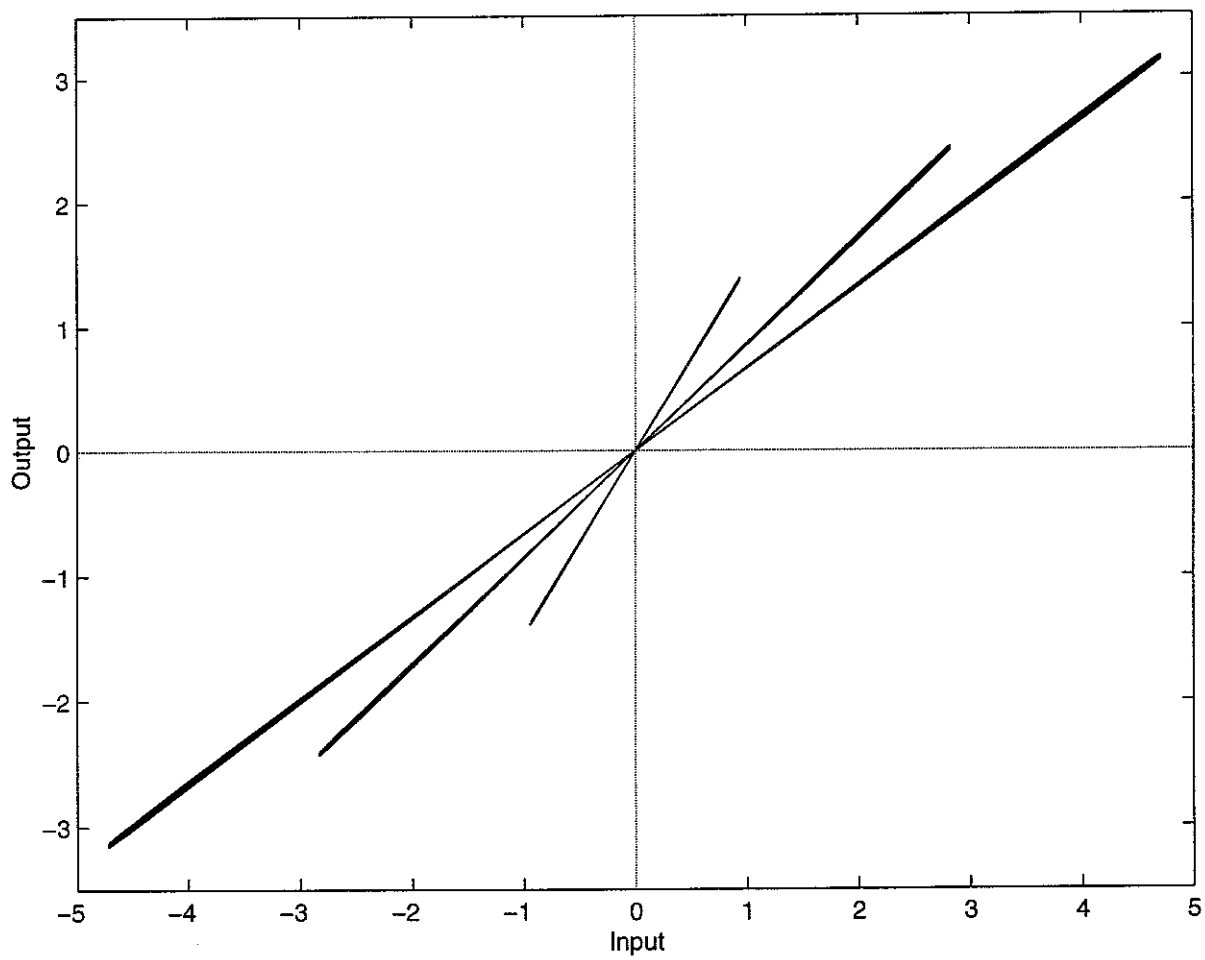

Figure 3.2: Instantaneous input-output characteristics of the automatic gain control system at three different input amplitudes.

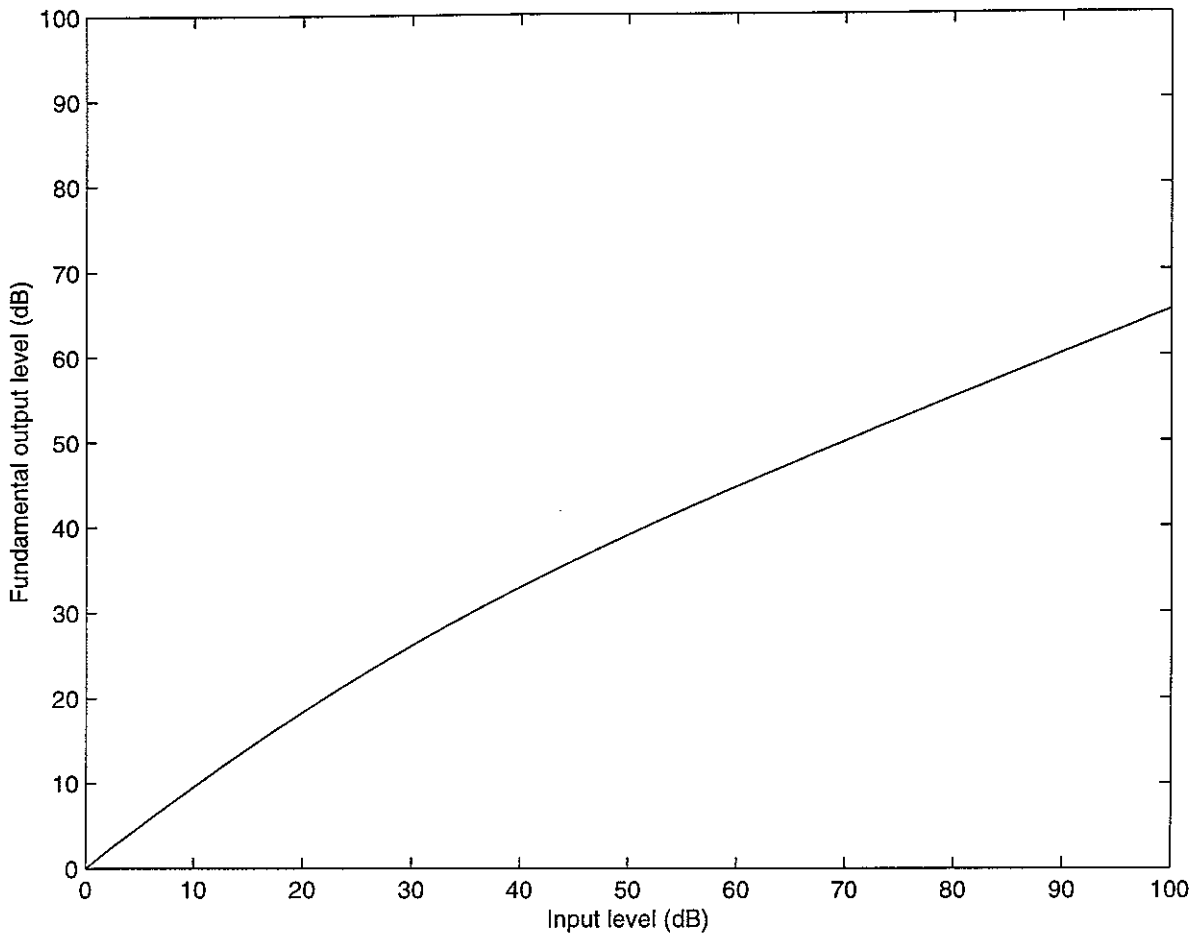

Figure 3.3: The level curve of the AGC with the gain law $g = (Y + \varepsilon)^{-1}$.

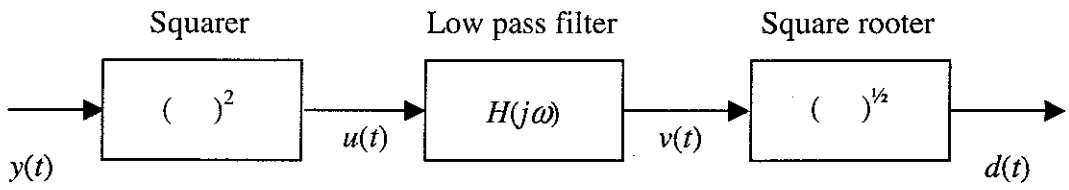

Figure 3.4: Assumed structure of the output detector in Figure 3.1, which estimates the rms amplitude of the output signal.

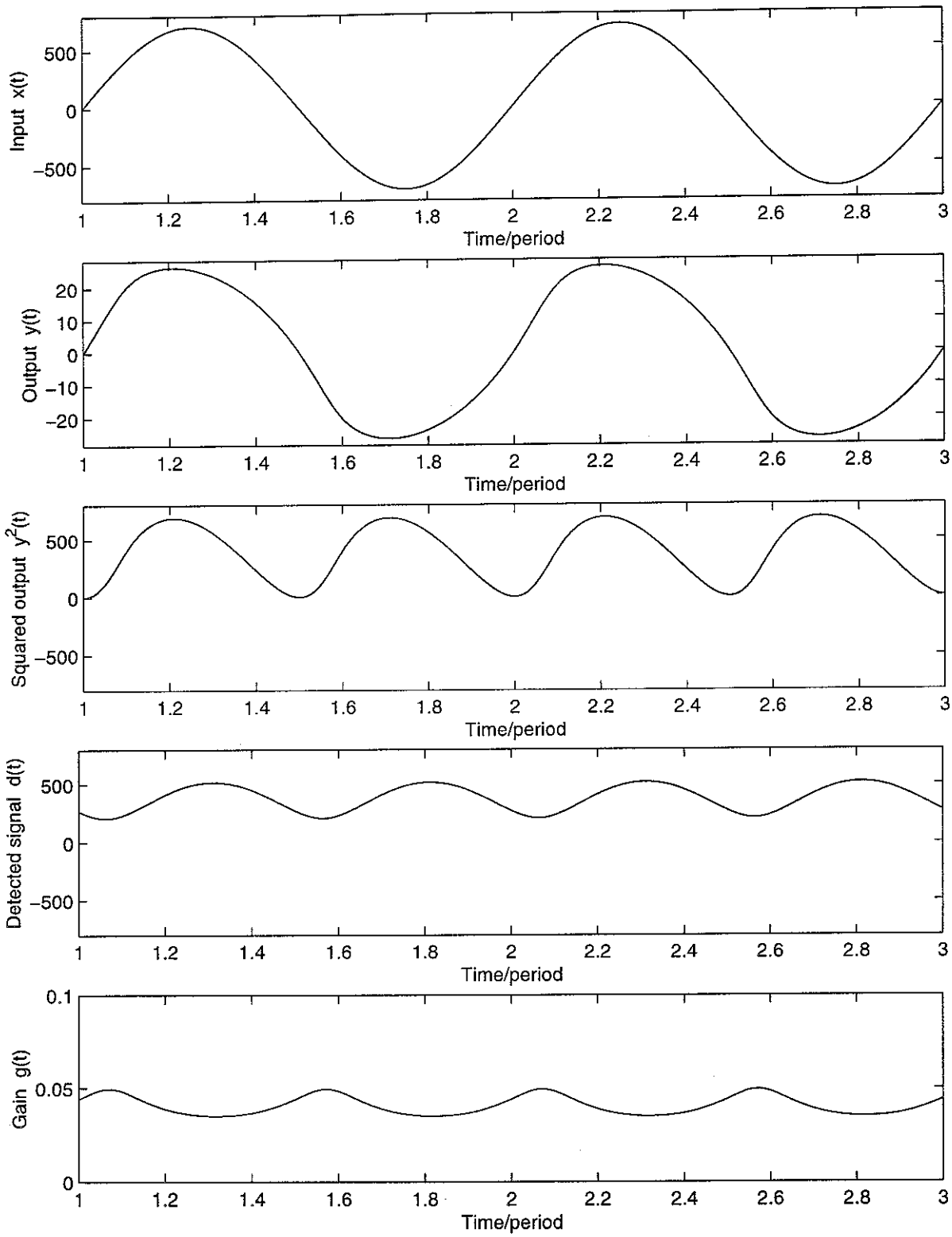

Figure 3.5 Waveforms in the AGC when the excitation frequency is equal to the cut-off frequency of the detector's low pass filter. From top to bottom: input signal, $x(t)$; output signal, $y(t)$; squared output signal, $y^2(t)$; detected signal, $d(t)$; and variation of gain with time, $g(t)$

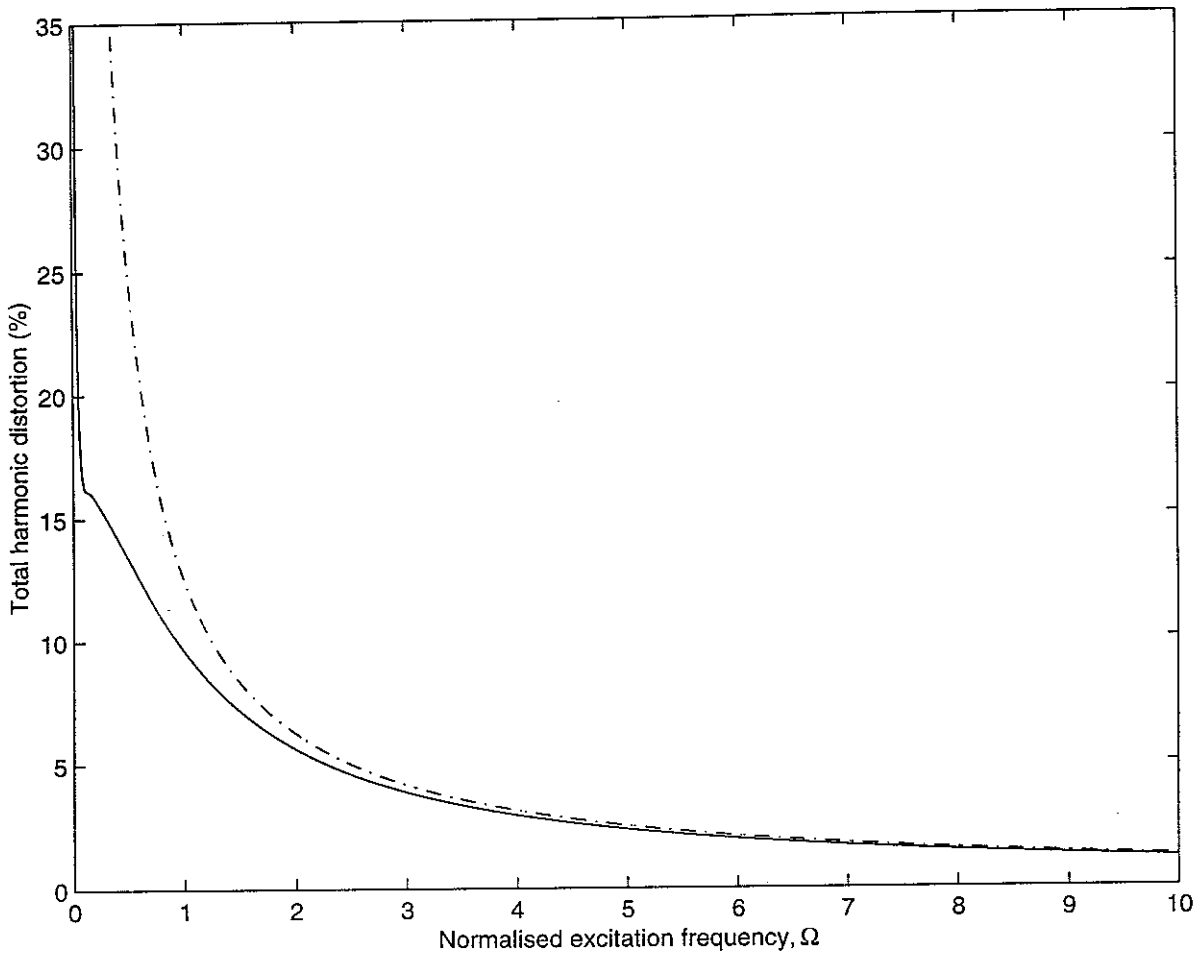

Figure 3.6: Total harmonic distortion of the AGC for a sinusoidal input signal as a function of the ratio of excitation frequency to the cut-off frequency of the low pass filter, Ω. The solid curve is the result of computer simulations and the dashed curve is the theoretical prediction, that THD = 1/8Ω.

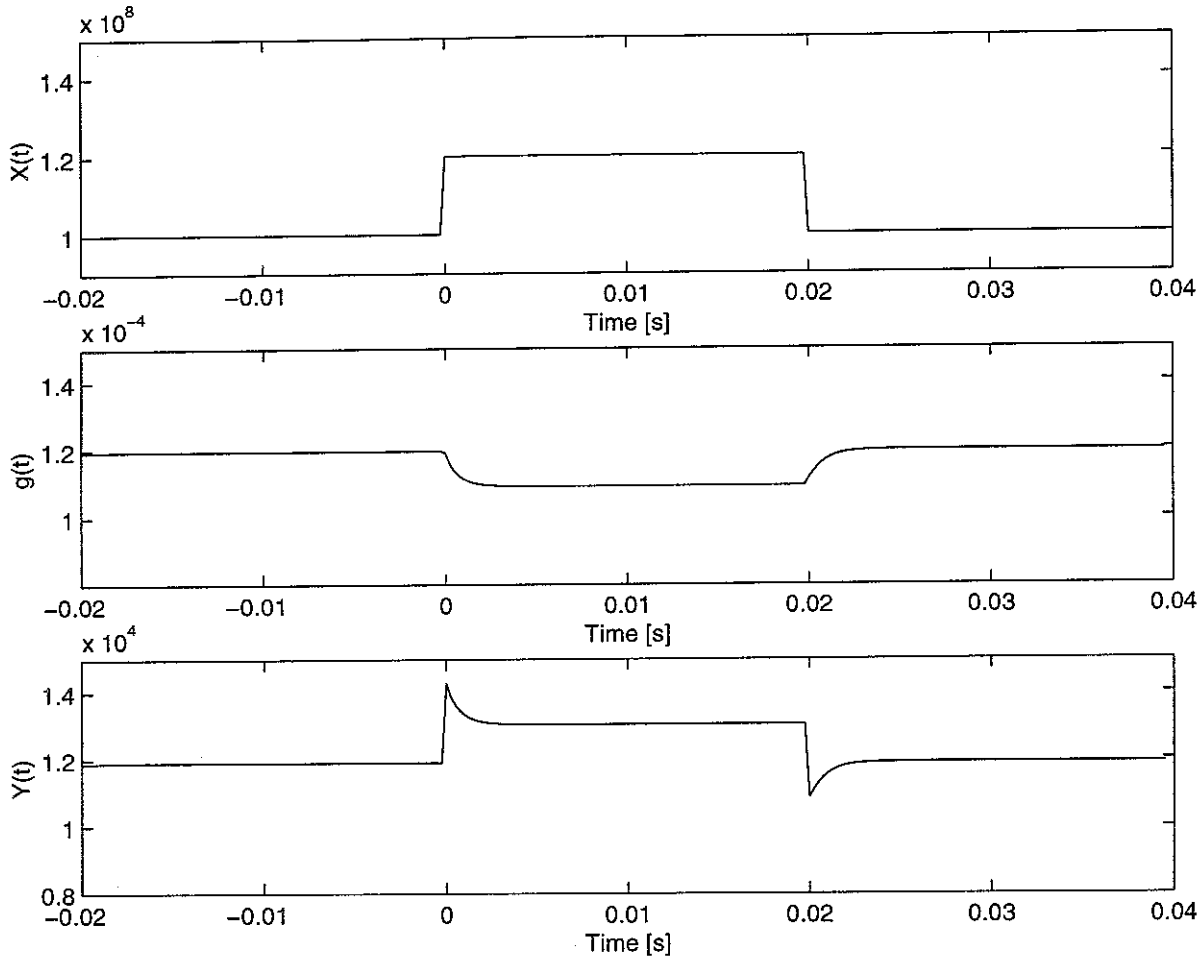

Figure 3.7 Transient response of the AGC to a sinusoidal input signal, whose amplitude is increased by 20% and then returned to its original value, showing the variation of input amplitude, $X(t)$; AGC gain, $g(t)$; and output amplitude, $Y(t)$

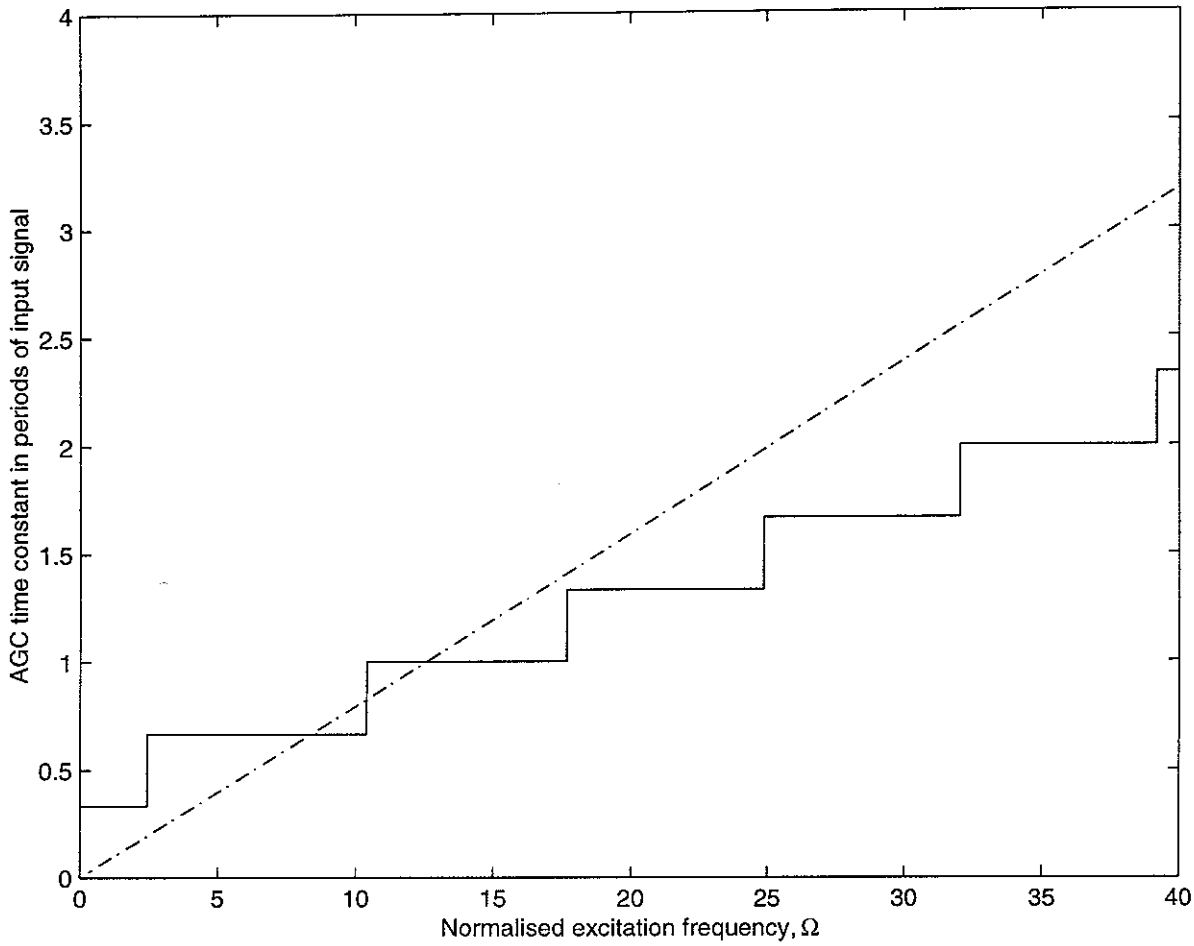

Figure 3.8: The time constant of the AGC to small perturbations, τ, in units of periods of the input signal, T_o, plotted against the excitation frequency divided by the cut-off frequency of the low pass filter, Ω, solid curve is the results of computer simulation and the dashed curve is the theoretical prediction, that $\tau/T_o = \Omega/4\pi$.

4. Dynamic Models with Amplitude-Dependent Damping

Individual parts of the basilar membrane are often modelled as dynamic, mass-spring-damper systems. The model is linear and passive if the damping term is constant and positive, but it is found that such passive models do not reproduce the sharp tuning observed in the healthy cochlea. In order to reproduce realistic tuning curves, a damping term which is negative over a certain frequency range must be used. The dynamic system can then generate as well as dissipate energy, and it becomes active. To reproduce the nonlinear properties of the cochlear response, however, the damping must be amplitude-dependent.

4.1 Van der Pol's Equation

A second-order dynamic models with amplitude-dependent damping can be represented in the form of a generalised Van der Pol equation (Guckenheimer and Holmes, 1983; Hanggi and Riseborough, 1983), whose differential equation may be written as

$$m\ddot{w} + c(w)\dot{w} + kw = f \tag{4.1}$$

where w is the waveform of the displacement response, $\dot{w}$ is the velocity and $\ddot{w}$ the acceleration, f is the waveform of the driving force, m and k are the constant mass and stiffness terms and $c(w)$ is the amplitude-dependent damping. The specific case

$$c(w) = c_o + c_1 w^2 , \tag{4.2}$$

with positive c_o, describes the passive nonlinear model of Hall (1974) and Talmadge et al. (1998).

If c_o is negative in equation (4.2), the system is active and nonlinear, and can display a wide range of dynamic behaviour, as described for example by Guckenheimer and Holmes (1983), including limit cycle oscillation, complicated entrainment, or synchronisation, behaviour when driven near its natural frequency and a variety of bifurcations, including Hopf bifurcations. A bifurcation is an abrupt change in the behaviour of a system caused by the variation of a control parameter. The limit cycle

behaviour of a Van der Pol oscillator, in particular, has been used to describe several features of spontaneous otoacoustic emissions (SOAE), including phase locking and synchronisation, suppression by external tones, the statistical fluctuations of the amplitude and phase and relaxation dynamics (as noted, for example, by de Boer, 1991, and Murphy et al., 1995).

Guckenheimer and Holmes (1983) describe a method of analysing the response of the Van der Pol equation involving a transformation of variables and averaging of the response. This reduces the second order nonlinear equation (4.1) to two coupled first order nonlinear equations, which can then be expressed as a single first order nonlinear equation in a complex variable. This method can be used to normalise and transform equations (4.1) and (4.2) to give the model equation used by Equíluz et al. (2000), described in the next section.

4.2 *Level Characteristic of the Hopf Bifurcation Model*

Equíluz et al. (2000) and Camalet et al. (2000) have used the tools of dynamic systems theory to explain a number of the nonlinear features found in hearing. Specifically they considered the compression of dynamic range, the sharper cochlear tuning for softer sounds and the generation of combination tones, and demonstrate that all of these features could be reproduced using a mathematical model of the type described above, which can exhibit a Hopf bifurcation. There is no output signal from an undriven system exhibiting a Hopf bifurcation if the control parameter is below a threshold value. Above this value of the control parameter the output is that of a limit cycle oscillation, whose amplitude rises as the control parameter is increased.

Equíluz et al. (2000) considered a generic equation describing a Hopf bifurcation as a model for cochlear response. Their model equation when driven by a complex excitation $x(t)$ may be written as

$$x(t) = \dot{y}(t) + \left(\eta + |y(t)|^2\right) y(t) - j\omega_n y(t), \tag{4.3}$$

where $y(t)$ is now the complex response of the system, ω_n is its natural frequency and η is a damping parameter (which is equal to $-\mu$ if μ is Equíluz et al.'s "control parameter").

If the damping parameter η is positive, then the system is stable and $y(t) = 0$ if $x(t) = 0$. If η is negative, however, then even with no input the system spontaneously oscillates. The oscillation amplitude is limited by the $|y(t)|^2 y(t)$ term in equation (4.3), which increases the damping as the response amplitude increases until the system displays a limit cycle whose amplitude is approximately $\eta^{1/2}$. It is the variation of the undriven amplitude with the parameter η that characterises the system as having a Hopf bifurcation.

When the system in equation (4.3) is driven by a complex periodic excitation of the form $x(t) = X e^{j\omega t}$, and it is assumed that the output is locked onto the excitation frequency, so that $y(t) = Y e^{j(\omega t+\phi)}$, then the level curve for this system may be derived. By considering the modulus squared amplitude of the response, its amplitude, Y, can be shown to be related to the forcing amplitude X by the equation

$$X^2 = Y^6 + 2\eta Y^4 + \left[\eta^2 + (\omega - \omega_n)^2\right] Y^2. \tag{4.4}$$

When the system is driven at its natural frequency, $\omega = \omega_n$, and the system is set at exactly its bifurcation point, i.e. $\eta = 0$, then the response amplitude is given (Equíluz et al., 2000) by

$$Y = X^{1/3}. \tag{4.5}$$

The input-output level curve for this system under these conditions thus has a constant slope of ⅓ dB/dB.

We now consider a slightly modified form of equation (4.3), in which the damping increases in proportion to $|y(t)|\, y(t)$ instead of $|y(t)|^2 y(t)$, so that

$$x(t) = \dot{y}(t) + \left(\eta + |y(t)|\right) y(t) - j\,\omega_n\, y(t). \tag{4.6}$$

In this case the locked response to a complex sinusoidal forcing of amplitude X has an amplitude Y given by

$$X^2 = Y^4 + 2\eta Y^3 + \left(\eta^2 + (\omega - \omega_n)^2\right) Y^2. \tag{4.7}$$

When the system is at its bifurcation point, $\eta = 0$, and is driven at its natural frequency, so that $\omega = \omega_n$, then the response amplitude in this case is equal to

$$Y = X^{1/2}, \tag{4.8}$$

and the level curve for the modified system has a slope of ½ dB/dB. Moreover if we assume that the system is set slightly away from bifurcation point, so that η is small, but is still driven at its natural frequency, then equation (4.7) can be written as

$$X^2 = Y^2\left(Y^2 + 2Y\eta + \eta^2\right), \tag{4.9}$$

so that

$$X = Y(Y + \eta). \tag{4.10}$$

If the system is below the bifurcation point, so that $\eta > 0$, and the forcing level is very small, then the response is also small. Provided $Y << \eta$, then equation (4.9) predicts that

$$Y \approx \frac{X}{\eta} \tag{4.11}$$

i.e. the response is linear with a 1 dB/dB level curve. As the driving level is increased, the response will increase until $Y > \eta$, in which case Y is approximately proportional to

$X^{½}$ again, and the slope of the level curve decreases to ½ dB/dB. Figure 4.1 shows the level curve for the locked solution to equation (4.6) when it is set slightly below the bifurcation point, $\eta > 0$, as discussed above, at the bifurcation point, $\eta = 0$, in which case equation (4.8) applies, and slightly above the bifurcation point, $\eta < 0$.

When the system is above the bifurcation point, $\eta < 0$, the system has a locked response with a ½ dB/dB slope for larger inputs, but for small inputs the output does not fall off with input level, but maintains a constant output level even if the input is removed. This corresponds to a limit cycle oscillation, which is reminiscent of the spontaneous otoacoustic emission. If η is negative then equation (4.10) is satisfied even if the system is undriven, $X = 0$, in which case the response level is given by $Y = -\eta$, which is a positive number. This simple model of the cochlear response is thus attractive because it can reproduce many of the features of real otoacoustic emissions, including spontaneous emissions at frequencies corresponding to certain positions on the cochlear where the effective low-amplitude damping is negative.

When the system is below the bifurcation point, $\eta > 0$, the level curve has exactly the same form as the AGC discussed in Section 3. In particular when the AGC has the modified gain law given by equation (3.9), the relationship between the forcing amplitude and the response amplitude is given by equation (3.10), which is of exactly the same form as equation (4.10) if $\eta > 0$.

Thus if the variation of damping with instantaneous response in the original Van der Pol's equation is selected correctly, we can obtain an identical level curve to that of the model where the gain depends on the average output level. It would be difficult to distinguish between these models in terms of their sinusoidal input output characteristics.

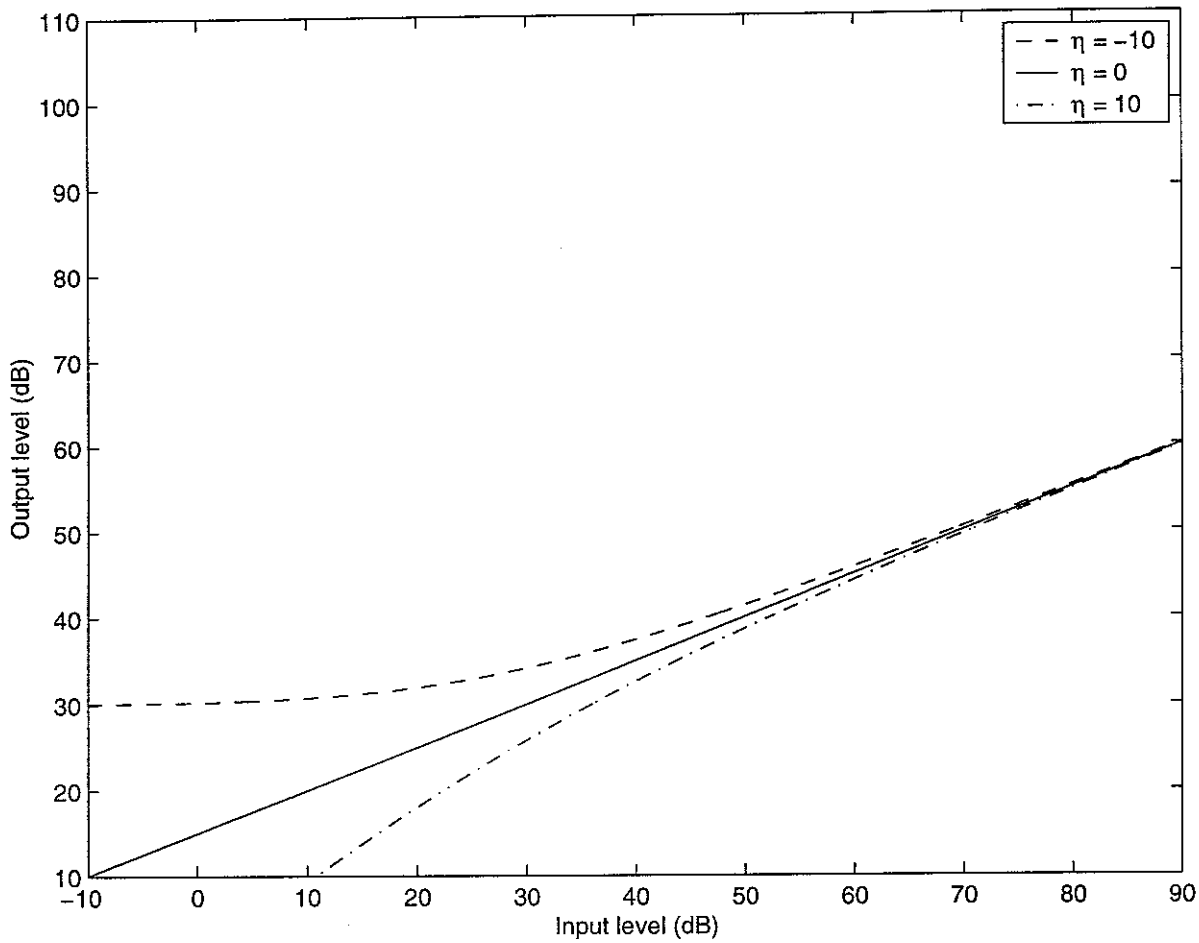

Figure 4.1: Level curve of the modified Hopf-bifurcation equation (4.6) when driven at its natural frequency with its low level damping, either positive (dot-dash curve), zero (solid curve) or negative (dashed curve).

5. Summary and Conclusions

Several different models of BM response have been considered. Of particular concern has been the response of these models to tonal excitation at a variety of input amplitudes. The level of the fundamental component of the output response plotted against the level of the input signal gives the level curve of the system, which is widely used to characterise the compressive nonlinearity of the BM response.

The first model of BM response considered is an instantaneous power-law nonlinearity. It is shown that a fixed power law gives a level curve with a single slope. By defining an instantaneous nonlinearity with a power law that varies with the instantaneous input amplitude, however, a level curve can be produced that has the required slope of 1 dB/dB at low levels and ½ dB/dB at higher levels.

The second model considered is an automatic gain control, whose gain is dependent on the output amplitude (a feedback AGC). With a suitable choice of gain law, a level curve can again be produced with a slope of 1 dB/dB at low levels and ½ dB/dB at higher levels. The distortion properties and transient properties of such an AGC are also analysed, and it is shown that it is possible to design such an AGC that responds within a cycle of the input waveform, and yet has a lower harmonic distortion than the instantaneous nonlinearity that produces an equivalent level curve.

One possible mechanism whereby the gain of a system could be varied by the output amplitude is for the system to have a nonlinear characteristic, whose bias position depends on the output level, as discussed in Appendix D.

A Van der Pol equation model for BM response is then discussed, which leads to the Hopf bifurcation model of Equíluz et al. (2000). In this model the damping depends in a nonlinear way on the instantaneous value of the output waveform. This is shown to have a steady-state gain that depends on the level of the output signal, and thus behaves in a similar way to a level-dependent nonlinear system. It is also shown how, with suitable modification of the parameters, the model can give a level curve with a slope of 1 dB/dB at low levels and ½ dB/dB at higher levels. Another example of a dynamic

nonlinear model for BM response that generates a similar form of level curve is the nonlinear feedback model of Yates (1990), which is discussed in Appendix E.

The Hopf bifurcation model is interesting because it clearly shows how a damping-like control parameter can be varied to produce either a stable system with the level characteristic described above, or a system which also has a ½ dB/dB slope in its level curve at higher amplitudes, but for low level excitation can produce limit cycle oscillations that are very reminiscent of spontaneous otoacoustic emissions.

Figure 5.1 shows the level curves computed from the final versions of the instantaneous nonlinearity in Section 2, and the level-dependent system in Section 3. These are asymptotically very similar and it may be difficult to choose between these two competing models on the basis of these steady-state level curves.

Figure 5.2 shows the transient response, to a suddenly applied tone, of this instantaneous nonlinearity and the level-dependent system consisting of a feedback AGC with a detector cut-off frequency which is half the input frequency. Although the level-dependent system takes some time to respond to the input, this time is much less than a cycle of the input waveform for such a well-adjusted AGC. It may thus be difficult to distinguish between these two types of nonlinear system in terms of their transient response, because the response time of the level-dependent system can be made so short that it looks almost instantaneous. Both systems also exhibit harmonic distortion in the steady state.

Figure 5.3 shows the instantaneous value of the output waveform plotted against the instantaneous value of the input waveform, for a tonal excitation at three different amplitudes, in the case of the two principal types of nonlinear system considered above. These three input-output characteristics overlay one another for the instantaneous nonlinearity, and thus form a single-valued nonlinear input-output characteristic. The instantaneous input-output characteristics of the AGC, however, are straight lines when the input frequency is well above the cut-off frequency of the output level detector, although the slope of these lines, reflecting the gain of the AGC, change with the amplitude of the input signal.

Of the three measures of response shown in Figures 5.1, 5.2 and 5.3, it would thus appear that the instantaneous input-output characteristic may be the clearest method of distinguishing between whether the cochlear nonlinearity is better modelled as an instantaneous or a level-dependent nonlinear system.

It has been shown that each of the models of cochlear response considered here can produce level curves that are consistent with the measured BM response data. The instantaneous nonlinear model has an input-output characteristic which is single-valued, however, whereas the level-dependent nonlinear model has an instantaneous input-output characteristic that is multiple-valued. As well as providing a possible method of distinguishing between the response of the two models, this observation has the important theoretical implication that a level-dependent system cannot be characterised by a Volterra series. The Wiener series and higher-order frequency responses thus also do not provide a complete description of such systems, despite their prevalence as theoretical models in hearing research.

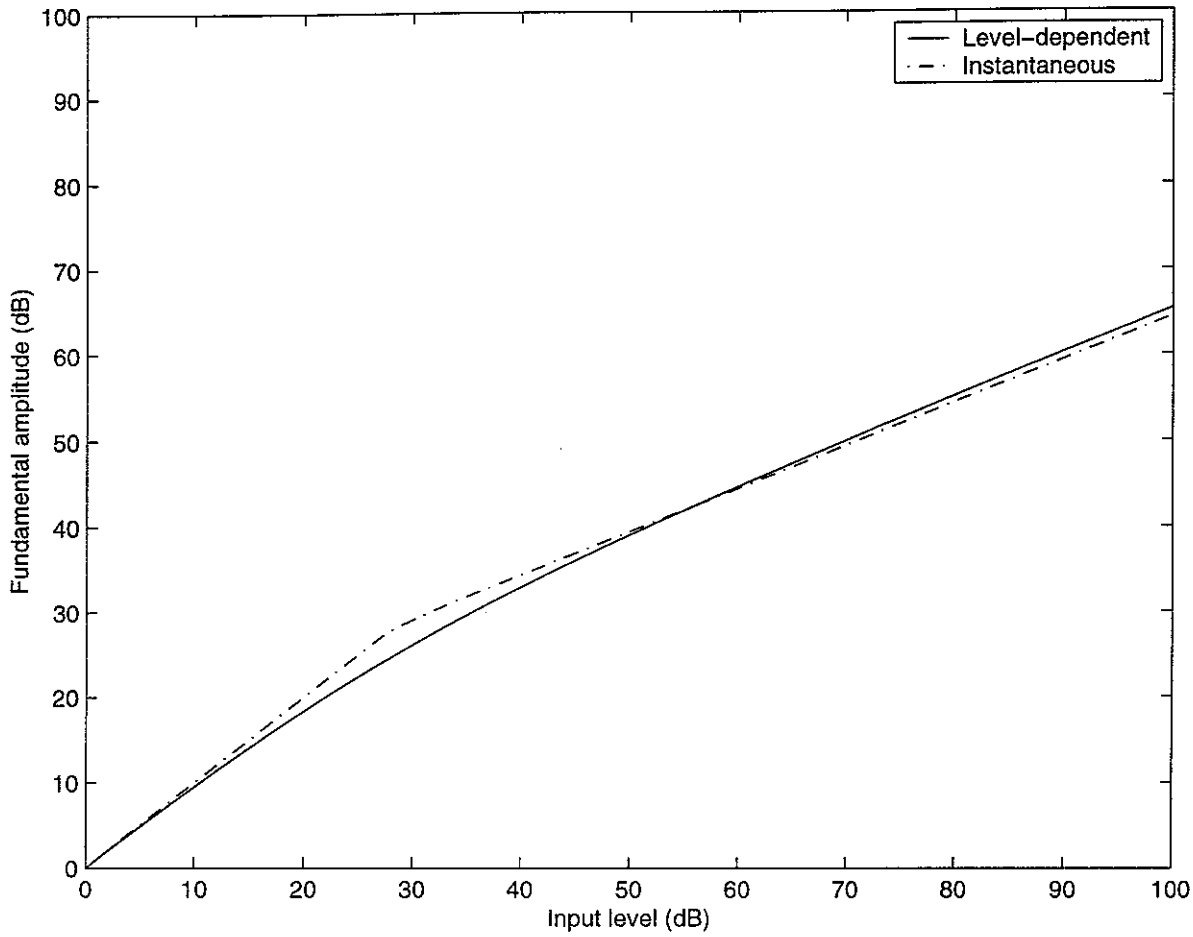

Figure 5.1 The level curve for sinusoidal excitation of the level-dependent nonlinear system, feedback AGC (solid line), and instantaneous nonlinear system, with scheduled power law (dashed line).

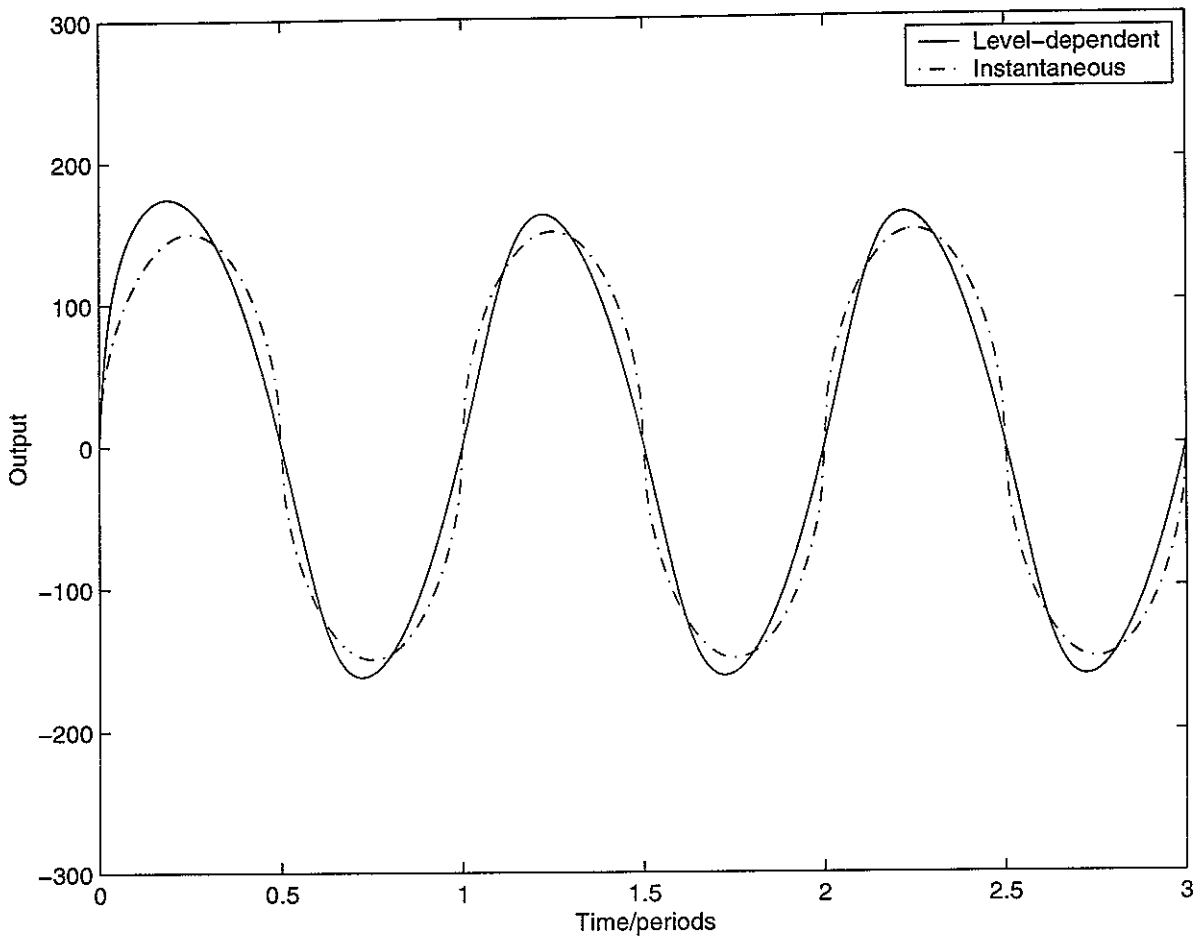

Figure 5.2 Output waveform, for an abruptly-applied sinusoid input, for the level-dependent nonlinear system, feedback AGC with cut-off frequency set equal to half the input frequency, solid line and instantaneous nonlinear system, with scheduled power law (dashed line).

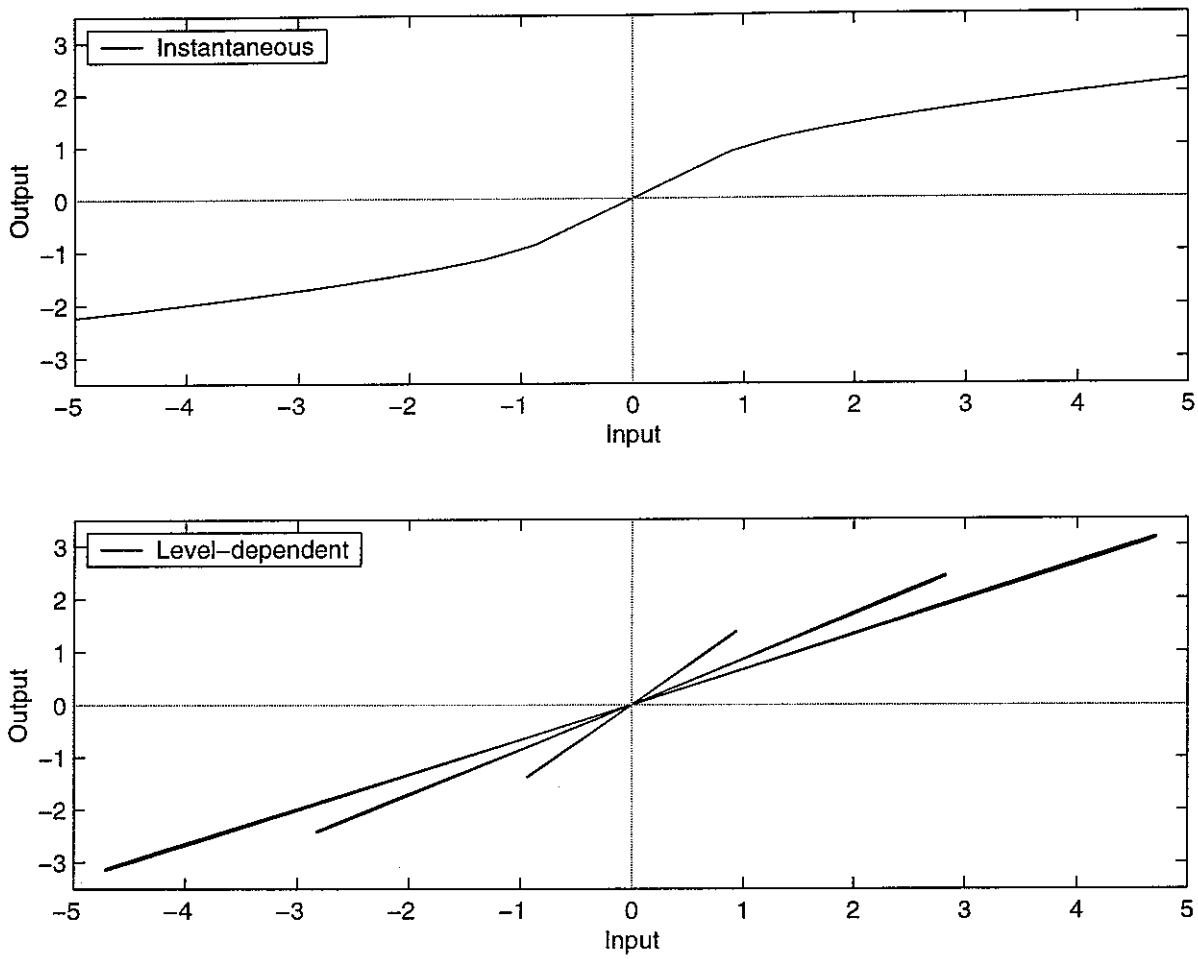

Figure 5.3 Instantaneous input-output characteristic for sinusoidal excitation at three levels of the instantaneous nonlinear system, with scheduled power law (upper) and the level-dependent nonlinear system, feedback AGC (lower).

APPENDIX A

NARMA Description of a Feedback AGC System

Assume that the feedback automatic gain control (AGC) is defined by the block diagrams of Figures 3.1 and 3.4, but that it is implemented in discrete time so that its input-output relationship is defined by

$$y(n) = g(n)\,x(n). \tag{A.1}$$

The gain law is assumed to be

$$g(n) = \frac{g_o}{d(n)+\varepsilon}, \tag{A.2}$$

in which g_o and ε are constants, and $d(n)$ is the output of the detector

$$d(n) = \left(v(n)\right)^{\frac{1}{2}}, \tag{A.3}$$

where $v(n)$ is the output of the low pass filter averaging the squared output of the AGC, given by

$$v(n) = (1-\delta)\,v(n-1) + \delta\; y^2(n-1), \tag{A.4}$$

in which δ is a constant related to the averaging time of the low pass filter. Note that the low pass filter has an input of $y^2(n-1)$ instead of $y^2(n)$ to ensure that the AGC loop can be causally implemented. Equations (A.1), (A.2) and (A.3) may together be written as

$$y(n) = \left[\left(v(n)\right)^{\frac{1}{2}} + \varepsilon\right]^{-1} g_o\, x(n). \tag{A.5}$$

Also, assuming $y(0) = 0$ and $v(0) = 0$ as initial conditions, the output of the low pass filter can be expressed as

$$v(n) = \sum_{i=1}^{n} (1-\delta)^{i-1} \delta\, y^2(n-i). \tag{A.6}$$

Substituting (A.6) into (A.5) gives an explicit, but nonlinear difference equation for the output of the AGC in terms of its current input and previous outputs:

$$y(n) = \left[\left(\sum_{i-1}^{n} (1-\delta)^{i-1} \delta\, y^2(n-i) \right)^{\frac{1}{2}} + \varepsilon \right]^{-1} g_o\, x(n). \tag{A.7}$$

APPENDIX B

Harmonic Distortion of the Feedback AGC

An estimate of the total harmonic distortion of the feedback AGC can be obtained by iteratively solving the equations describing the detector for a sinusoidal input. Initially, we assume that the output signal is undistorted so that

$$y(t) = \sqrt{2}\, Y \cos(\omega_o t), \tag{B.1}$$

where for the compressive AGC with a gain law given by equation (3.8), $Y = \sqrt{X}$. The output of the squaring device in Figure 3.4 is then

$$y^2(t) = Y^2[1 + \cos(2\omega_o t)], \tag{B.2}$$

which is also the input to the low pass filter, *u(t)*. The low pass filter will pass the d.c. component of $y^2(t)$ with a gain of unity, but the second-harmonic component will be affected by the complex gain

$$H(j2\omega_o) = M\, e^{j\theta}. \tag{B.3}$$

The output of the low pass filter is thus

$$v(t) = Y^2[1 + M\cos(2\omega_o t + \theta)]. \tag{B.4}$$

The overall output of the detector, which determines the gain, is thus

$$d(t) = Y[1 + M\cos(2\omega_o t + \theta)]^{\frac{1}{2}}, \tag{B.5}$$

and the time varying gain of the AGC, assuming a simple compressive gain law, is

$$g(t) = \frac{1}{d(t)} = Y^{-1}\left[1 + M\cos\left(2\omega_o t + \theta\right)\right]^{-\frac{1}{2}}. \tag{B.6}$$

Assuming the second harmonic component of the low pass filter output is not too large, so that $M < 1$, then

$$g(t) \approx Y^{-1}\left[1 - \frac{M}{2}\cos\left(2\omega_o t + \theta\right)\right]. \tag{B.7}$$

The output waveform can be written as

$$y(t) = g(t)\,x(t), \tag{B.8}$$

and since $x(t) = \sqrt{2}\,X\cos\left(\omega_o t\right)$ and using (B.7) then

$$y(t) \approx \sqrt{2}\,X\,Y^{-1}\cos\left(\omega_o t\right)\left[1 - \frac{M}{2}\cos\left(2\omega_o t + \theta\right)\right]. \tag{B.9}$$

Assuming $Y \approx \sqrt{X}$, then the output waveform is

$$y(t) \approx \sqrt{2X}\left[\cos\left(\omega_o t\right) - \frac{M}{2}\left[\cos\left(\omega_o t\right)\cos\left(2\omega_o t + \theta\right)\right]\right], \tag{B.10}$$

which can be written as

$$y(t) \approx \sqrt{2X}\left[\cos\left(\omega_o t\right) - \frac{M}{4}\left[\cos\left(\omega_o t + \theta\right) + \cos\left(3\omega_o t + \theta\right)\right]\right]. \tag{B.11}$$

Assuming that M/4 << 1, then the fundamental component of the output has an rms amplitude $Y = \sqrt{X}$, as assumed above, but the output also now has a third harmonic component of rms amplitude $\sqrt{X}\,\frac{M}{4}$. The total harmonic distortion of the AGC is thus approximately

$$\text{THD}\,(\%) \approx \frac{M}{4} \times 100\% \,. \tag{B.12}$$

For the first order low pass filter described in Section 3 above, then

$$M = \left|H(j2\omega_o)\right| = \left|\frac{\omega_L}{\omega_L + j2\omega_o}\right| \tag{B.13}$$

and further assuming that $2\omega_o >> \omega_L$, then

$$M \approx \frac{\omega_L}{2\omega_o} = \frac{1}{2\Omega} \tag{B.14}$$

where Ω is the ratio of the excitation frequency to the cut-off frequency of the low pass filter. A simple approximation for the total harmonic distortion in such an AGC is thus

$$\text{THD}(\%) \approx \frac{1}{8\Omega} \times 100\% \tag{B.15}$$

APPENDIX C

Time Constant of a Feedback AGC

The nonstationary tonal input to a feedback automatic gain control (AGC) system can be written

$$x(t) = \sqrt{2}\, X(t)\cos(\omega_o t), \tag{C.1}$$

where $X(t)$ is an envelope function. We assume that $X(t)$ has a value of X_1 until $t = 0$ and then abruptly changes to X_2

$$X(t) = \begin{cases} X_1 & \text{for } t > 0 \\ X_2 & \text{for } t \geq 0 \end{cases}. \tag{C.2}$$

It is also assumed that the input change is not too large, so that $X_1 - X_2 << X_1$. The output of the AGC is also assumed to be tonal but nonstationary and can be written in the form

$$y(t) = \sqrt{2}\, Y(t)\cos(\omega_o t), \tag{C.3}$$

where $Y(t)$ is the envelope of the output, which is itself equal to

$$Y(t) = g(t)\, X(t), \tag{C.4}$$

where $g(t)$ is the instantaneous gain of the AGC.

The AGC is assumed to be compressive with the steady-state gain law $g = 1/Y$ so that $Y = \sqrt{X}$ and so the steady state gain can also be written as

$$g = \frac{1}{\sqrt{X}}. \tag{C.5}$$

Let

$$g_1 = \frac{1}{\sqrt{X_1}}, \tag{C.6}$$

which is the gain of the system for $t < 0$, and

$$g_2 = \frac{1}{\sqrt{X_2}} \tag{C.7}$$

which is the steady state gain as $t \rightarrow \infty$.

We now make the working assumption that the instantaneous gain changes exponentially from g_1 to g_2 at $t = 0$, i.e.

$$g(t) = g_2 + e^{-t/\tau}\left(g_1 - g_2\right) \tag{C.8}$$

where τ is the overall time constant of the AGC. The assumed form of the instantaneous gain, equation (C.8), can also be written as

$$g(t) = g_2\left(1 + \Delta e^{-t/\tau}\right), \tag{C.9}$$

where Δ is the fractional change in gain

$$\Delta = \frac{g_1 - g_2}{g_2}, \tag{C.10}$$

which is assumed to be much less than unity since the change in the input level is small.

The envelope of the output is thus

$$Y(t) = g(t)\,X_2 \qquad \text{for } t > 0 \tag{C.11}$$

so that, using equation (C.9),

$$Y(t) = g_2 X_2\left(1 + \Delta e^{-t/\tau}\right). \tag{C.12}$$

The output signal is squared in the feedback loop of the AGC to give

$$y^2(t) = 2\,Y^2(t)\cos^2\left(\omega_o t\right), \tag{C.13}$$

which is equal to

$$y^2(t) = Y^2(t)\left[1 + \cos\left(2\omega_o t\right)\right], \tag{C.14}$$

and it is assumed that the alternating component is smoothed out by the low pass filter, but that the squared envelope function still excites its dynamics. Assuming the low pass filter is a first order system, with input $u(t)$ and output $v(t)$, its dynamics are defined, as in Section 3, by the differential equation

$$u(t) = v(t) + T_L \frac{dv(t)}{dt}\ , \tag{C.15}$$

where T_L is the time constant of the low pass filter. Within the AGC, the low pass filter input is the squared output signal, $y(t)$, whose dynamics on the timescales of interest here are determined by its envelope function, so that

$$u(t) = Y^2(t) = {g_2}^2\,{X_2}^2\left(1 + 2\Delta e^{-t/\tau}\right) \tag{C.16}$$

where equation (C.12) and the fact that $\Delta << 1$ have been used to obtain the final expression.

The output of the low pass filter is square-rooted and used to define the gain of the AGC, so that

$$g(t) = \frac{1}{\sqrt{v(t)}}. \tag{C.17}$$

Thus

$$v(t) = \left(g(t)\right)^{-2}\ , \tag{C.18}$$

and using the assumed form of $g(t)$ in equation (C.9), then the input to the low pass filter can be written as

$$v(t) = \left(g_2\left[1+\Delta e^{-t/\tau}\right]\right)^{-2}, \tag{C.19}$$

and again assuming that Δ is small, then

$$v(t) = \frac{1}{g_2{}^2}\left[1-2\Delta e^{-t/\tau}\right]. \tag{C.20}$$

The right hand side of equation (C.15) may thus be written

$$v(t)+T_L\frac{dv(t)}{dt} = \frac{1}{g_2{}^2}\left[1+2\left(\frac{T_L}{\tau}-1\right)\Delta e^{-t/\tau}\right]. \tag{C.21}$$

Using equation (C.7) we note that

$$g_2{}^2 X_2{}^2 = \frac{1}{g_2{}^2}, \tag{C.22}$$

so that equation (C.16), which is the left hand side of equation (C.15), is equal to equation (C.21), which is the right hand side of equation (C.15), provided

$$\frac{\tau_L}{\tau}-1=1. \tag{C.23}$$

The assumptions we have made about the form of the gain variation with time, equation (C.8), is thus consistent with all the equations for the AGC provided equation (C.23) is

satisfied. The time constant of the gain change, τ, is thus given by the solution to equation (C.23), which is

$$\tau = \frac{T_L}{2} \tag{C.24}$$

where T_L is the time constant of the low pass filter.

APPENDIX D

The Generation of a Level-Dependent Gain by Biasing a Nonlinear Function

Figure D.1 shows one possible mechanism by which a level-dependent gain with the characteristics discussed in Section 3 may be produced without explicit gain-scheduling. In this arrangement the rms level of the alternating component of the output is used to bias a nonlinear function. The a.c. gain of the nonlinear function is dependent on its slope, which changes with the bias point. Such a mechanism would be consistent with the otherwise puzzling observation, noted by de Boer (1991), that the d.c. displacement of the BM change with the amplitude of an alternating excitation.

In Figure D.1 the input to the nonlinear function, $u(t)$, is the sum of the alternating input signal;

$$x(t) = \sqrt{2}X\cos(\omega t)\,, \tag{D.1}$$

and a d.c. bias term, which is assumed to be dependent on the rms level of the fundamental component of the output signal, $z(t) = Y_1$, with an offset ε, so that

$$u(t) = Y_1 + \varepsilon + x(t)\,. \tag{D.2}$$

The output waveform will also have a d.c. bias, Y_o, and harmonic components Y_n, and may be written as

$$y(t) = Y_o + \sqrt{2}Y_1\cos(\omega t) + \sqrt{2}Y_2\cos 2(\omega t) + \cdots, \text{etc.}\,. \tag{D.3}$$

It is further assumed that the nonlinear function is logarithmic, so that

$$y(t) = \ln\left(u(t)\right) \tag{D.4}$$

where ln denotes a natural logarithm. Substituting equation (D.2) into equation (D.4) gives

$$y(t) = \ln\left(Y_1 + \varepsilon + x(t)\right) \tag{D.5}$$

so that

$$y(t) = \ln\left(Y_1 + \varepsilon\right) + \ln\left(1 + \frac{\mathrm{x(t)}}{\mathrm{Y_1} + \varepsilon}\right). \tag{D.6}$$

Assuming that $|x(t)|/(Y_1 + \varepsilon) < 1$, the final term in equation (D.5) can be expanded as a series to give

$$y(t) = \ln\left(Y_1 + \varepsilon\right) + \frac{x(t)}{Y_1 + \varepsilon} + \frac{x^2(t)}{2\left(Y_1 + \varepsilon\right)^2} + \cdots . \tag{D.7}$$

If $x(t)$ is sinusoidal, as in equation (D.1), then substituting this into equation (D.7) and comparing each term in equation (D.3), we can see that

$$Y_o \approx \ln\left(\mathrm{Y_1} + \varepsilon\right), \tag{D.8}$$

$$Y_1 \approx \frac{X_1}{Y_1 + \varepsilon}\ , \tag{D.9}$$

and

$$Y_2 \approx \frac{{X_1}^2}{4\left(Y_1 + \varepsilon\right)^2} = \frac{{Y_1}^2}{4}\ . \tag{D.10}$$

The term Y_o is the d.c. bias on the output, which is proportional to the log of the output amplitude plus a bias term. The fundamental component of the output has exactly the same form as equation (3.10), and so the corresponding level curve has a slope of 1dB/dB for values of Y_1 less than about ε, and a slope of ½dB/dB for values above ε. The main source of output distortion is predicted to be the second harmonic in this case, and the total harmonic distortion is approximately

$$\mathrm{THD}(\%) \approx \frac{Y_2}{Y_1} \times 100\% = \frac{Y_1}{4} \times 100\% \ , \qquad \text{(D.11)}$$

which rises linearly with output amplitude.

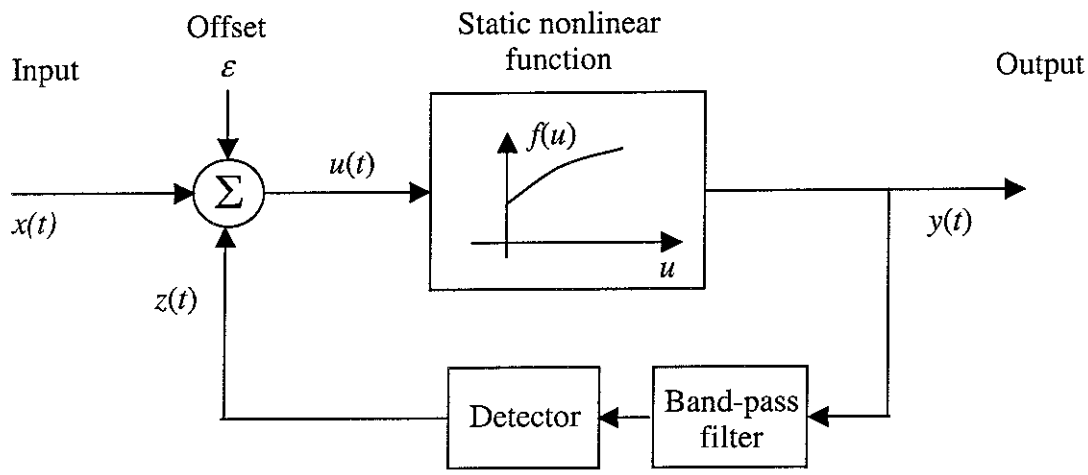

Figure D.1: A possible mechanism for a level-dependent gain using a nonlinear function whose bias point is set by the output amplitude.

APPENDIX E

A Nonlinear Positive Feedback System

Figure E.1 shows the system with a nonlinear feedback path proposed by Yates (1990) as a model for the basilar membrane input-output function, following similar systems described by Zwicker (1979). In this figure G is a gain, Φ is a nonlinear network, β is a frequency-selective network that provides a gain, also denoted by β, at the fundamental frequency of the input signal, which is assumed to be tonal, but attenuates its harmonics.

The nonlinear network denoted Φ in Figure E.1 is assumed by Yates (1990) to have an instantaneous input/out characteristic

$$v(t) = \frac{A\, y(t)}{A + y(t)}, \tag{E.1}$$

where $y(t)$ is the input waveform to Φ, $v(t)$ is its output waveform and A is a threshold amplitude. The instantaneous gain characteristic can thus be expressed as

$$\Phi = \frac{v(t)}{y(t)} = \frac{A}{A + y(t)}. \tag{E.2}$$

The nonlinear network behaves linearly at low input amplitudes, i.e.

$$\Phi \approx 1, \text{ so that } v(t) \approx y(t) \text{ if } |y(t)| < A \tag{E.3}$$

and saturates at high amplitudes, i.e.

$$\Phi \approx \frac{A}{y(t)}, \text{ so that } v(t) \approx A \text{ if } |y(t)| > A. \tag{E.4}$$

The instantaneous input-output characteristics of this nonlinear network and its gain characteristic are plotted in Figure E.2 for $A = 10^4$, which is the value typically used by Yates

(1990). Only the upper right quadrant of the characteristic is plotted, since a logarithmic amplitude scale is used, but the characteristic is the same for negative values of v and y since it is symmetric. The gain of Φ is unity at very low values of input and progressively drops as the input value rises.

The ratio of the complex output, Y, of the overall nonlinear feedback system shown in Figure E.1 at the driving frequency to the complex input, X, is taken by Yates (1990), to be

$$\frac{Y}{X} = G_{\text{closed}} = \frac{G}{1-\beta\Phi G}, \tag{E.5}$$

where G and β are real constants, and Φ is taken to be a real number that depends on the amplitude of the output signal.

For the configuration shown in Figure E.1, however, the nonlinear system, Φ, is assumed to act instantaneously and $y(t)$ in equation (E.1) is a time-varying waveform. The gain of Φ, given in equation (E.2) will thus vary above and below unity during one cycle of a sinusoidal variation of Y, and the output waveform from Φ will be distorted. The network β is then assumed to select only the fundamental component of $v(t)$. The entire feedback network could thus be characterised by a describing function, whose form can be computed using the methods described, for example, by Ogata (1970). Alternatively, equation (E.2), with $y(t) = Y$, could itself be taken to define the describing function of the feedback network, as is implicitly assumed by Yates (1990) in deriving equation (E.5).

In this case the gain of the complete feedback path at the excitation frequency can be written as

$$\beta\Phi = \frac{\beta A}{A+Y}. \tag{E.6}$$

Since there is no phase shift through G or through the feedback path, all of the potentially complex signals, X, Y and V can be taken to be real in this case, and the overall transfer function at the driving frequency, equation (E.5), has no phase shift.

It should be noted that Figure E.1 is a *positive* feedback system and is only stable if the loop gain, $\beta \Phi G$, is less than unity. Yates (1990) assumes that G is always unity and that β is typically 0.999. For very low level signals, such that Φ is linear and has a gain of one, $\beta \Phi G$ is very close to unity (0.999) and the system is thus very close to instability. The closed loop gain in equation (E.5) is large under these conditions, 10^3, but is also extremely sensitive to small changes in gain around the feedback loop. Such a change would be caused by an increase in the output signal, so that the gain of Φ drops below unity. If, for example, the value of Y in equation (E.6) is just 1% of A, then the magnitude of Φ becomes about 0.99, and the closed loop gain drops from 10^3 to about 10^2, i.e. by about 20dB.

If the gain of the overall feedback path in Figure E.1 is denoted by H, then equation (E.6) indicates that for tonal excitation, the magnitude of H is effectively being scheduled on the amplitude of the feedback loop's output, Y, as shown in Figure E.3(a). The system shown in Figure E.3 is a level-dependent nonlinear system. The use of the describing function has allowed the instantaneous nonlinearity to be replaced by a level-dependent nonlinearity only for the particular case of a steady-state tonal excitation at the centre frequency of the network β. The dynamics of the output detector in Figure E.3(a) do not affect its steady state response, but would affect its transient behaviour. Thus Figure E.3(a) cannot be taken to be a representation of the behaviour of the system in Figure E.1 for any other excitation signal than a steady state tone.

Since G and H are entirely real in Figure E.3(a), the response of the whole feedback loop can be replaced by a real number representing its closed loop gain, g, where

$$g = \frac{G}{1 - GH}, \tag{E.7}$$

which is implicitly scheduled on the output amplitude because of the dependence of H on this amplitude in Figure E.3(a). The resulting block diagram, Figure E.3(b), is exactly the same as that of the automatic gain control (AGC) considered above. In this case, however, the gain law is slightly different to that considered previously.

Assuming, as in Yates (1990), that $G = 1$ and that H is given by equation (E.6), then the dependence of the gain on the amplitude of the output amplitude may be written as:

$$g = \frac{1}{1-\beta\Phi} = \frac{A+Y}{(1-\beta)A+Y}. \tag{E.8}$$

If we assume that $A >> Y$, but that $(1-\beta)$ is correspondingly small, so that $(1-\beta)A$ is not too large and can be written as ε, then the gain law is equal to

$$g = \frac{1}{Y+\varepsilon}, \tag{E.9}$$

exactly as in Section 3 above. In general, however, equation (E.8) can be written as

$$g = \frac{Y+A}{Y+\varepsilon}, \tag{E.10}$$

where $\varepsilon = (1-\beta)A$ again and $(1-\beta) < 1$ so that $A > \varepsilon$.

The behaviour of the system described by equation (E.10) can be split into the three distinct regions described by Yates (1990), but using the formulation above, analytic approximations can now be derived for the characteristics in each region. First, for very low levels:

$$\text{if } Y << A \text{ and } Y << \varepsilon \quad , \quad \text{then } g \approx \frac{A}{\varepsilon}, \tag{E.11}$$

so that the system behaves linearly, with a level curve having a slope of 1dB/dB, and a gain of $A/\varepsilon > 1$. Second, for moderate levels:

$$\text{if } Y < A \text{ but } Y > \varepsilon, \text{ then } g \approx \frac{A}{Y}, \tag{E.12}$$

so that the system behaves in a compressive way, as described in Section 3 above, with a level curve slope of ½dB/dB. Finally, for very high input levels:

$$\text{if } Y > A \text{ and } Y > \varepsilon \text{ then } g \approx 1, \tag{E.13}$$

so that the system again behaves linearly, with a level curve slope of 1dB/dB, but with a significantly lower gain than at very low levels. Figure E.4 illustrates the level curve for an AGC with the gain low given by equation (E.10) for $\varepsilon = 10^3$ and $A = 10^9$.

It should be noted that the frequency-selective network and nonlinear network do not have to be in the feedback path, as shown in Figure E.1 to produce an output-dependent level curve. If, for example, these elements were positioned in the feedforward path and the feedback path was assumed to have a unity gain, then the closed loop gain becomes

$$g = \frac{\beta \Phi}{1 - \beta \Phi} = \frac{\beta A}{Y + \varepsilon} \tag{E.14}$$

where equation (E.6) has been used for $\beta\Phi$ and ε is again equal to $(1-\beta)A$. The gain law in equation (E.14) does not revert to a linear one at high amplitudes and so is consistent with those derived in the sections above.

Finally, it is interesting to note that if β is chosen to be slightly greater than unity, instead of slightly less than unity and the frequency selective network is a resonant system, then the system will produce limit cycle oscillations for very low input, exactly like the Van der Pol oscillator.

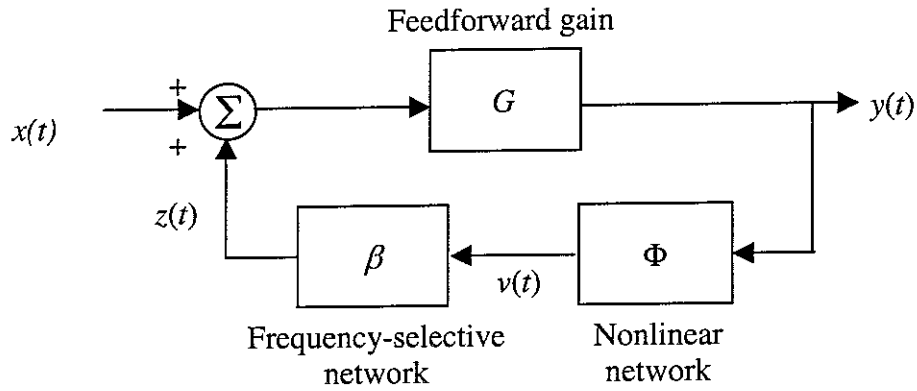

Figure E.1: Block diagram of a system with nonlinear feedback path.

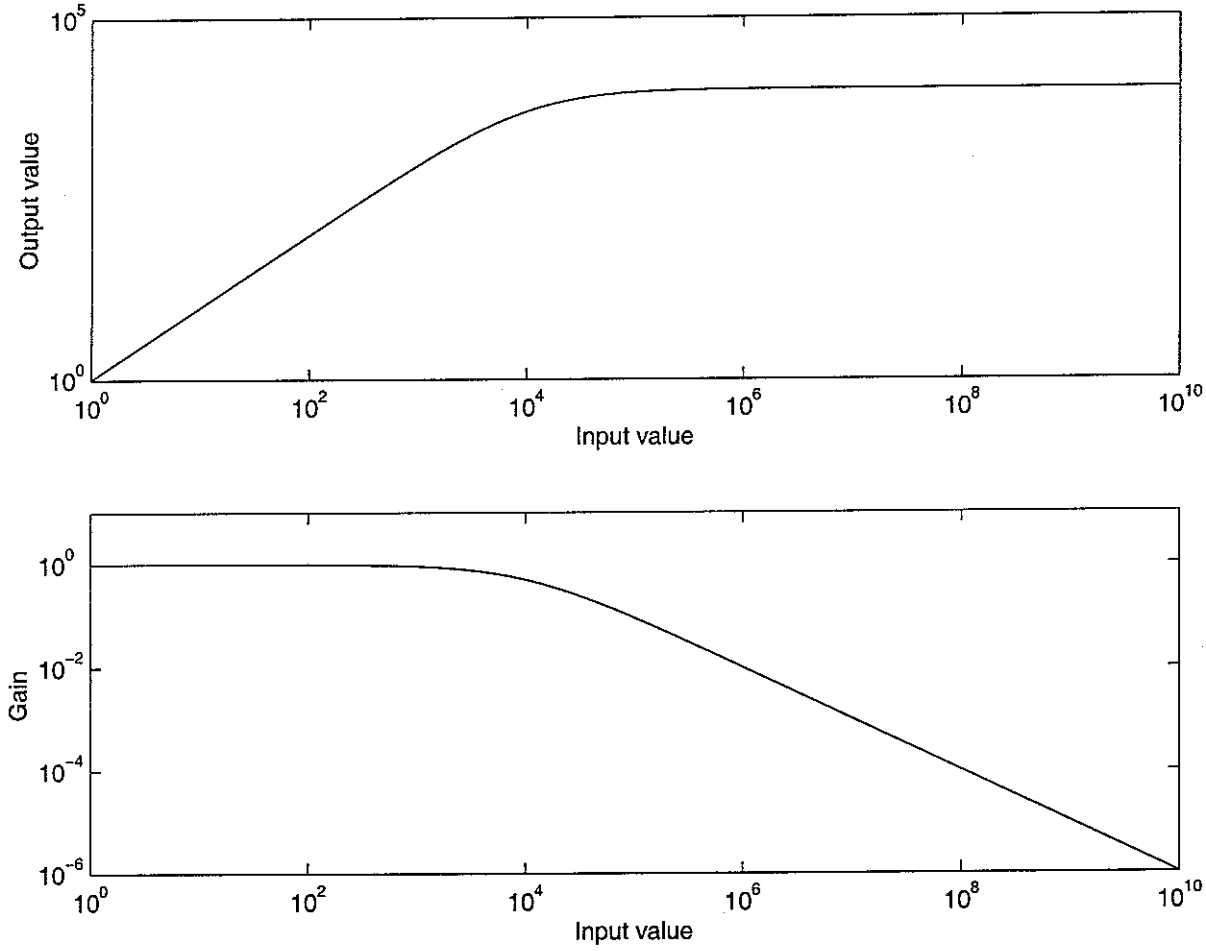

Figure E.2: The assumed input-output function (upper) and the gain characteristic (lower) for the nonlinear network Φ.

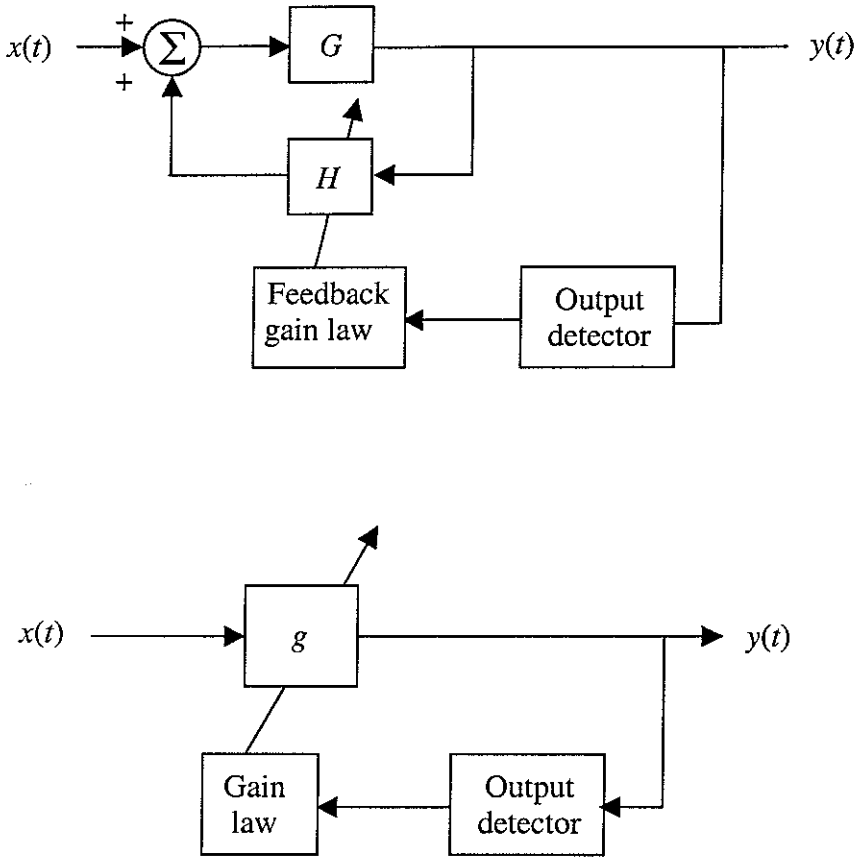

Figure E.3: Alternative interpretation of the equations describing a system with nonlinear feedback and frequency selectivity; as a feedback system whose feedback gain is scheduled on output lever (a), and the equipment automatic gain control if $g = G/(1 - GH)$ (b).

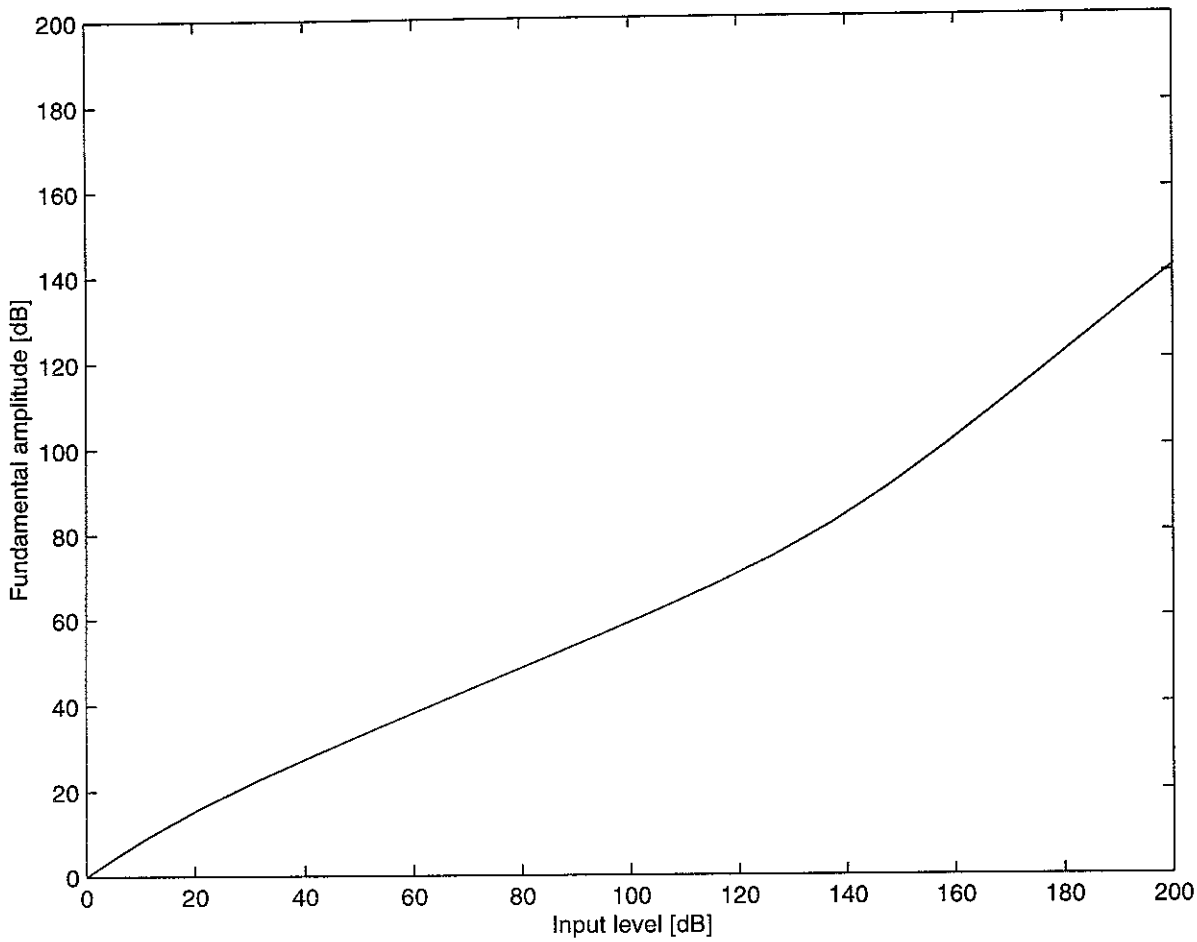

Figure E.4: The input-output level curve for an AGC having the gain law $g = (Y + A)/(Y + \varepsilon)$ where $A = 5\times10^{6}$ and $\varepsilon = 5\times10^{3}$.

References

Camalet S., Duke T., Jülicher F. and Prost J., 2000, Auditory sensitivity provided by self-tuned critical oscillations of hair cells. *Proc. Nat.Acad.Science*, **97**(2), 3183-3188.

de Boer E., 1991, Auditory physics. Physical principles in hearing theory. III. *Physics Reports*, **203**(3), 125-231.

Dorn P.A., Konrad-Martin D., Neely S.T., Keefe D.H., Cyr E. and Gorga M.P., 2001, Distortion product otoacoustic emission input/output functions in normal-hearing and hearing-impaired human ears. *Journal of the Acoustical Society of America*, **110**(6), 3119-3131.

Equíluz V.M., Ospeck M., Choe Y., Hudspeth A.J. and Magnasco M.O., 2000, Essential nonlinearities in hearing. *Physical Review Letters*, **84**(22), 5232-5235.

Goldstein J.L., 1995, Relations among compression, suppression and combination tones in mechanical responses of the basilar membrane: data and MBPNL model. *Hearing Research*, 89, 52-68.

Engebretson A.M. and Eldredge D.H., 1968, Model for the nonlinear characteristics of cochlear potentials. *Journal of the Acoustical Society of America*, **44**(2), 548-554.

Guckenheimer J. and Holmes P., 1983, *Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields. Springer-Verlaag.*

Hall J.L., 1974, Two-tone distortion products in a nonlinear model of the basilar membrane. *Journal of the Acoustical Society of America*, **56**(6), 1818-1828.

Hanggi P. and Riseborough P., 1982, Dynamics of nonlinear dissipative oscillators. *American Journal of Physics*, **51**(4), 347-352.

Kanis L.J. and de Boer E., 1993, Self-suppression in a locally active nonlinear model of the cochlea: A quasilinear approach. *Journal of the Acoustical Society of America*, **94**(6), 3199-3206.

Kim D.O., Molnar C.E. and Pfeiffer R.R.., 1984, A system of nonlinear differential equations modeling basilar-membrane motion. *Journal of the Acoustical Society of America*, **54**(6), 1517-1529.

Lyon R.F. and Mead C., 1988, An analog electronic cochlea. *IEEE Trans. On Acoustics, Speech and Signal Processing*, **36**(7), 1119-1133.

Lyon R.F., 1990, Automatic gain control in cochlear mechanics. *Proc. Mechanics and Biophysics of Hearing Conference, Wisconsin*, 395-402.

Meddis R. and O'Mard L.P., 2001, A computational algorithm for computing nonlinear auditory frequency selectivity. *Journal of the Acoustical Society of America*, **109**(6), 2852-2861.

Murphy W.J., Talmadge C.L., Tubis A. and Long G.R., 1995, Relaxation dynamics of spontaneous otoacoustic emissions perturbed by external tones. *Journal of the Acoustical Society of America*, **97**, 3702-3710.

Norton S.J. and Neely S.T., 1987, Tone-evoked otoacoustic emissions. *Journal of the Acoustical Society of America*, **81**(6), 1860-1872.

Nuttall A.L. and Dolan D.F., 1996, Steady-state sinusoidal velocity responses of the basilar membrane in guinea pig. *Journal of the Acoustical Society of America*, **99**(3), 1556-1565.

Ogata K., 1970, *Modern Control Engineering*. Prentice Hall.

Pfeiffer R.R., 1970, A model for two-tone inhibition of single cochlear-nerve fibers. *Journal of the Acoustical Society of America*, **48**(6), 1373-1378.

Pickles J.O., 1988, *An Introduction to the Physiology of Hearing*. Academic Press, 2nd Edition.

Rhode W.S. and Recio A., 2000, Study of mechanical motions in the basal region of the chinchilla cochlea. *Journal of the Acoustical Society of America*, **107**(6), 3317-3332.

Sellick P.M., Patuzzi R. and Johnstone B.M., 1982, Measurement of basilar membrane motion in the guinea pig using the Mössbauer technique. *Journal of the Acoustical Society of America*, **72**(1), 131-141.

Smoorenburg G.F. 1972, Combination tones and their origin. *Journal of the Acoustical Society of America*, **52**, 615-632.

Talmadge C.L., Tubis A., Long G.R. and Piskorski P., 1998, Modeling otoacoustic emission and hearing threshold fine structures. *Journal of the Acoustical Society of America*, **104**(3), 1517-1543.

Yates G.K., 1990, Basilar membrane nonlinearity and its influence on auditory nerve rate-intensity functions. *Hearing Research*, **50**, 145-162.

Zwicker E. and Schloth E., 1984, Interrelation of different oto-acoustic emissions. *Journal of the Acoustical Society of America*, **75**(4), 1148-1154.