Application of DSP Methods to Sound Reproduction

David Peter Round

Thesis Submitted in Candidature for the Degree of

Doctor of Philosophy

September 1996

School of Electronic Engineering and Computer Science

University of Wales, Bangor

United Kingdom

Abstract

Three different applications of Digital Signal Processing (DSP) technology to audio reproduction have been examined: Feedback control of loudspeakers using the back EMF, correction for loudspeaker Doppler distortion and room acoustic control.

Using the back EMF generated by the motion of the voice coil of a loudspeaker to sense the motion of the cone, and therefore allow the use of closed loop control, would potentially allow more accurate low frequency reproduction. Previous attempts to build a practical system have floundered due to the presence of the voice coil's self inductance. The application of system identification techniques to gain the parameters required to deal with this problem is examined. The concept developed worked well under simulation.

The motion of a loudspeaker cone is such that when signals of more than one frequency are reproduced there will be an element of Doppler distortion. A filter was developed using DSP technology that was demonstrated to significantly reduce the levels of this type of distortion in a real test.

Room acoustic control using DSP to produce destructive interference for the reverberant sound energy at the listening location has been demonstrated to a degree in the past by a number of other researchers. In this document a feasibility study of the application of this technology to HiFi stereo systems is examined. In addition a number of possible modifications to the classic Normalized Least Mean Squared (NLMS) derived algorithm currently used for acoustic control, the "filtered-x" algorithm, are examined.

It was found that there are a severe practical limitations to the bandwidth over which it is possible to correct for deficiencies in the room response. The most serious problem is the inherent distance that separates the listener and the sensor in a real system. No acceptable solution to this problem is believed possible. The practical bandwidth over which a system of this type will operate usefully is approximately 100Hz.

Statement of Originality

The work presented in this thesis was carried out by the candidate, except where otherwise stated. It has not been presented previously for any degree, nor is it at present under consideration by any other degree awarding body.

Candidate: D. Round

David Peter Round

Statement of Availability

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organizations.

Candidate: D. Round

David Peter Round

Acknowledgments

I would like to express my sincere thanks to Dr. Mike Lorenz for his invaluable advice and guidance throughout the course of this research.

I would like to thank Berkeley Camera Engineering, California, for the loan of the digital audio interface and DSP hardware.

Dedication

To my wife, Tina, and family.

Contents

Chapter 1

Introduction

The quest for perfect sound reproduction has absorbed a great deal of research effort over the years and has resulted in some high quality (and very expensive) equipment. Many of the problems with sound reproduction lie with the frequency range and dynamic range of the human hearing system. The sound waves we can hear vary in length from greater than 17 metres down to something like 17 millimetres and while the dynamic range can be as large as 140dB [22], which is a ratio of 1 to 10^7. This leads to problems of linearity in loudspeaker design.

Recently the cost of Digital Signal Processing (DSP) hardware has fallen dramatically. It is now possible to realistically conceive of applying this technology to domestic HiFi reproduction. (In fact, given the continuing trend in the price of DSP hardware, there may be overall saving available if less stringent specifications of the loudspeakers can be traded for the application of this technology.)

Any signal can be approximated digitally by measuring its amplitude sufficiently regularly. There are benefits to recording a signal digitally, the most important is that it is now possible to copy and store the signal exactly. There is no further reason for further degradation to the signal quality once it has been converted into a digital representation. It is true that the digital version of the signal is an approximation, but if the measurements are sufficiently accurate and frequent then the loss of quality is not significant. Another advantage of having signals in digital form is that it is possible to process them with computer technology. Using techniques developed, very complicated filters can be built, far more complicated than would be practical using traditional analogue techniques. Processing signals this way has proved to be so useful that special computer hardware has been developed to run the programs more rapidly, so called Digital Signal Processing (DSP) hardware.

A number of possible applications of DSP hardware to the sound reproduction problem have been identified, namely, compensation for the room acoustics, feedback control of the cone motion and correction for Doppler effects in sound reproduction.

General improvements in state of the art analogue electronics and the introduction of digital storage media (CDs etc.) have improved the quality of reproduction, and highlighted the short-comings remaining electro-mechanical components.

There are two electro-mechanical components that are significant in a modern system, the microphone(s) and the loudspeaker(s). The former are only used for the initial recording and can therefore be precision devices since their cost is effectively spread over many listeners.

Loudspeakers are, however, required in every reproduction system. Loudspeakers also face a unique problem; in order to make a sufficiently loud sound particularly at low frequencies they are required to move a significant volume of air - which implies that the units need to be large. Consumer requirements and commercial interests require that (typically) the units need to be small and unobtrusive. Thus any way of in proving the performance of the small units is of particular interest.

In order to maintain the Sound Pressure Level (SPL) but reduce the size of the loudspeaker it is necessary to increase the travel of the cone. (The part of the loudspeaker that actually moves the air.) It is possible to design loudspeakers using current technology with larger maximum permissible cone excursions, but it becomes more difficult to maintain the accuracy of the output.

When attempting to improve the linearity of a system or extend the useful area of linear operation it is normal to make use of closed loop control techniques. These techniques appear attractive when attempting to improve the performance of loudspeakers over large cone excursions. Closed loop control takes a measure of the current system output and compares this with the desired response to the input signal. A correction is supplied on the basis of the difference between the two.

In order to apply closed loop techniques to the control of a moving coil loudspeaker some method is required to measure the movement of the cone. (Position, velocity and acceleration are all acceptable.) The moving part of the measurement hardware needs to be very light, since the performance of a loudspeaker is very much governed by the weight of the cone; inevitably for moving coil drivers the cone is significantly heavier than the air which it moves. The measurement method also needs to be at least as accurate as the existing system if it is not to impair performance. The most common method is to add an accelerometer but this inherently means adding mass to the system, not to mention the cost of the additional parts. The alternative considered here is to measure the back EMF, which is generated in a normal loudspeaker by the movement of the cone. There are problems that have prevented this idea form being used in the past but it is hoped that with the application of modern DSP technology these can be resolved.

An additional problem that is encountered as the cone size is decreased, and thus it is moving faster for the same SPL, is that distortions due to the Doppler effect become more significant. Correction for this type of distortion is difficult since it is non-linear and therefore linear filter DSP techniques will not work. The solution is to design a filter specific to this problem.

A slightly different problem is presented by the room used for the sound reproduction. The sound reproduced by the loudspeakers will be reflected off of the walls and thus the listener will hear not only the sound direct from the loudspeaker but these unplanned echoes. Since sound travels relatively slowly, about 340m/s, the delay between the reflected signal and the direct version is significant even in smaller rooms. Correction for this type of distortion is not easy due to the complexity of the echoes that will be received and is certainly only plausible using DSP technology.

Chapter 2

An Introduction to DSP Methods

2.1 Introduction

This chapter is a general introduction to digital filters and adaptive processes, considering in particular to their application to the specific problem of inverse filtering. The aspect of model identification will also be briefly considered, since this uses very similar technology.

Initially the rationale behind making use of digital filters is considered and the basic concepts are reviewed. Building on the very basics the underlying concepts of adaptive systems are discussed and some of the more common methods are considered in more detail.

Adaptive systems are an important aspect of modern control technology since they offer many advantages over the more traditional methods. Typically this type of system is able to find a good approximation to the ideal solution without human intervention. Such systems are especially desirable in the case of acoustics where the systems are complicated and unpredictable.

Adaptive filters are almost invariably digital since the parameters of a digital filter can easily be changed automatically during normal operation.

2.2 Basics

2.2.1 Analogue Filters

Traditionally all filters were analogue devices constructed typically of capacitors, resistors and perhaps inductors. (Inductors are avoided at audio frequencies since they are bulky and expensive.) Analogue filters can either be passive, using only capacitors, resistors and inductors, or active. Active filters typically use operational amplifiers to remove the need for inductors and therefore improve the performance and/or decrease the cost of the filter. Using these techniques it is possible to build a filter that may have several poles. Analogue filters are usually tuned once, at manufacture, by altering variable resistances etc.

This type of filter is usually designed by taking the desired frequency response and abstracting it into a realizable set of poles and zeros in the 's' domain. Well known configurations can then be used to generate the required pole-zero placements. It most circumstances it is sufficient to produce only the required amplitude response, which simplifies the problem, but most phase responses can be produced if the additional complication is accepted.

It is possible to make adaptive filters that adjust automatically by making use of passive components that alter their operational coefficients as a result of an applied electrical signal. Thus a capacitance may be replaced with a varicap, a device whose capacitance varies with applied voltage. Electrically variable resistances are also available, both in the form of an optically linked Light Dependent Resistance (LDR) and Light Emitting Diode (LED) and integrated devices. (A current mode multiplying DAC can also be used to produce effectively a digitally controlled variable resistance.) It is possible to construct electronically variable inductances but there are not commonly used. Typically electronically variable components suffer from even more pronounced parasitic effects then their fixed equivalents which makes the design of filters using them more difficult and reduces filter performance.

Even given that an electrically variable filter can be produced to perform a non-trivial task the design of the control gear and the choice of the measure that will be used to tune the parameters will be difficult. In practice recourse to digital techniques will almost certainly be required.

All simple analogue filters are inherently infinite impulse response types, there is no realistic equivalent to digital finite impulse response filters.

Given that the form of the system is known it is possible to construct inverse filters using analogue techniques. Since analogue delays are very limited and in flexible it is not be possible to produce an accurate, but delayed, inverse to a non-minimum phase system.

To summarize; the difficulties involved in designing analogue adaptive inverse filters are such that practical application of the technology is limited.

Switched capacitance filters are analogue filters but they operate in discrete time and thus offer some digital-like properties. This type of filter can be readily tuned over a significant but still limited range by altering the switch frequency. It is realistic to construct filters using this technique up to a few tens of poles.

2.2.2 Digital Filters

Digital filters offer many significant advantages over their analogue equivalents.

- No parasitic effects

  Digital filters do not suffer from the effects of parasitic imperfections that occur in real analogue components. There are imperfections in the response of digital filters that are related to computational errors etc. but these are deterministic.

- Changing parameters

  It is easy to change the parameters of a digital filter over a large range without degrading the performance of the filter, provided that sufficient computational power is available.

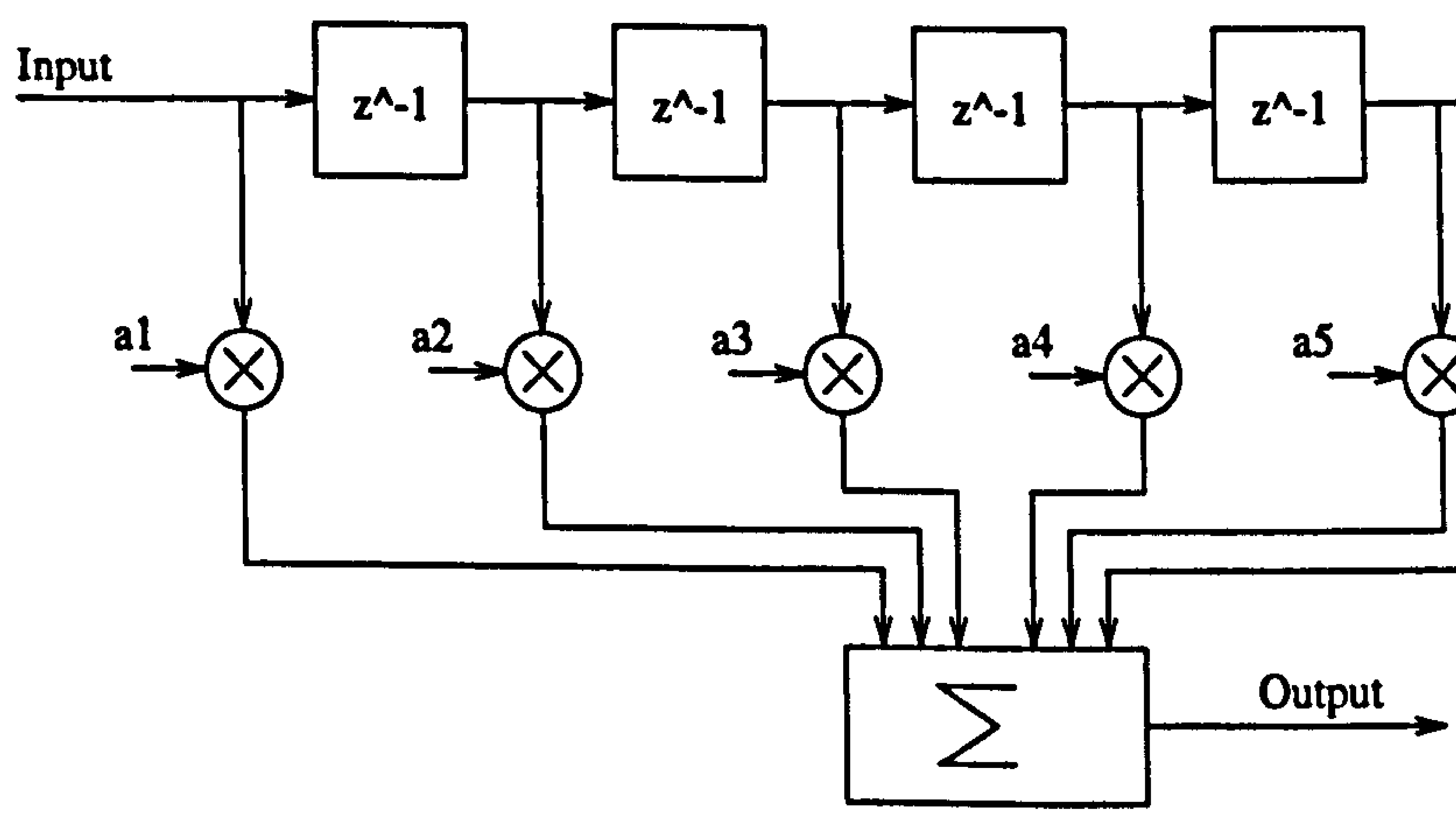
- Structure

  The structure of digital filters, particularly finite impulse response filters, makes automatic design easier to achieve.

  Digital filters can simulate any ideal analogue filter but in addition are capable of performing operations that would be difficult to implement using analogue technology.

There are two main classes of digital filters, Infinite Impulse Response (IIR) filters and Finite Impulse Response (FIR) filters. FIR filters are the simplest of the two, consisting of a delay line. The output is calculated as the sum of the input values that are stored in the delay-line after each is multiplied by fixed coefficients that define the filter. The concept is represented in the diagram below:

Figure 2.1: FIR Filter



The advantages of FIR filters over their IIR equivalents are:

- Unconditionally Stable
- Linear phase design possible

There are of course disadvantages:

- More processing required

  To achieve the same order more elements are required in an FIR filter.

- Delays

  Typically there are quite substantial delays associated with FIR filters.

IIR filters are the other main class of filters and these differ from FIR filters in that they operate by feedback. The output signal is fed back to the stages of the delay line. The main advantage of these filters is that it is sometimes possible to achieve a similar result to a long FIR filter with fewer coefficients. There are some significant disadvantages:

- More difficult design

  Adaptive processes are available for calculating IIR filters but they tend to be less robust than those for FIR filters. It is more difficult to design an IIR filter because filters of this type are not inherently stable. Thus it is necessary to ensure that all filters that could result from an automatic design process will be stable. The stability of this type of filter can be affected by computational errors because of their inherent recursion.

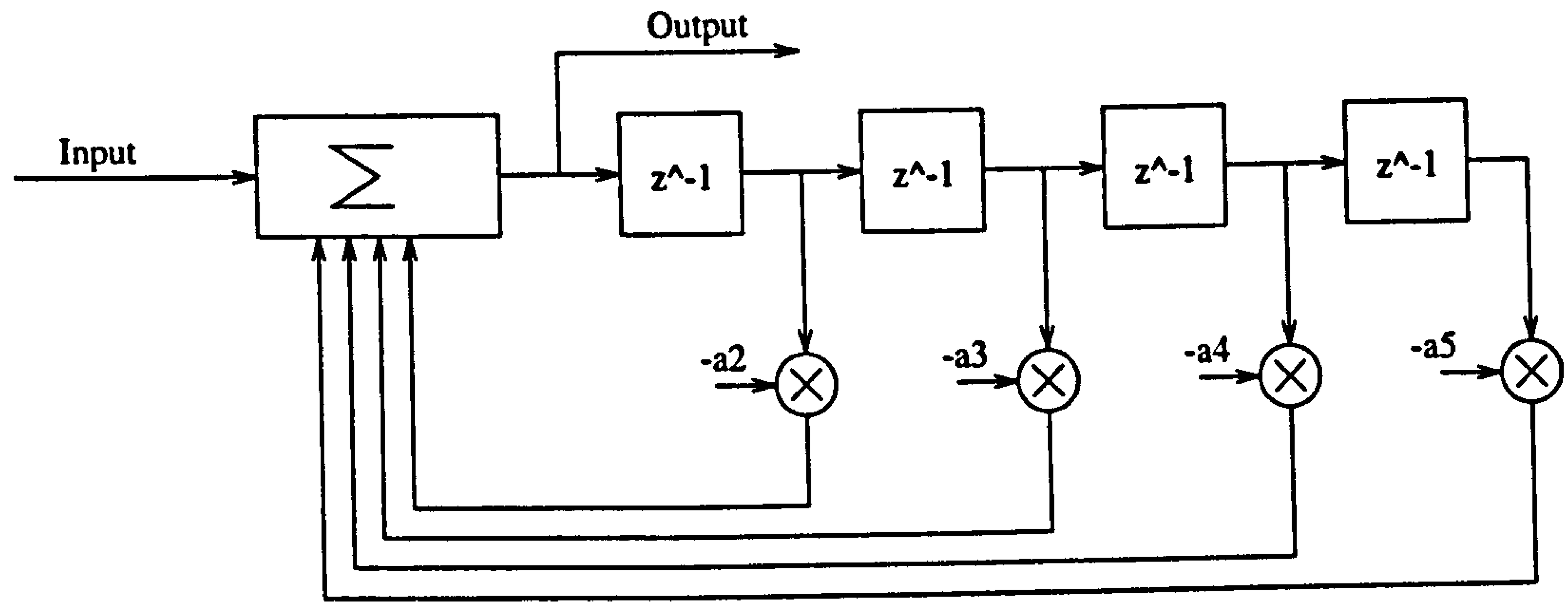
- Phase

  Phase distortion for this type of filter is usually highly non-linear.

The disadvantages are sufficient to make this type of filter impractical for this application.

A typical block diagram for an IIR filter is given in figure 2.2.

Figure 2.2: IIR Filter



2.3 Adaptive Filters

An adaptive system uses a performance measure to change its operational parameters and thus improve the performance of the system. The performance measure can be thought of as a surface in the number of operational parameters plus one dimensions - the 'extra' dimension being used for the performance measure itself. The performance surface can be complicated, with local minima, but complicated surfaces are difficult to search. Ideally the performance measure should be expressed so that the surface takes as simple a form as possible - in practice this means that the performance surface is typically quadratic, ie. the surface has a smooth gradient leading to a single minimum.

In this section methods for searching quadratic performance surfaces in particular are considered. It is acknowledged that not all performance surfaces are quadratic but these more complicated surfaces are outside the scope of this introduction. In the first instance methods for finding the minimum in two dimensions will be considered. (When searching for the optimum weights for an FIR filter each weight will represent a distinct dimension and there will be one dimension in addition to this, the performance measure. Thus initially the situation where there is only one weight plus the performance measure, ie. two dimensions will be considered.)

Normally the aim of adaptive system is to minimize some error function. Typically the magnitude of the error is used so that the situations where there are many dimensions can be accommodated.

2.3.1 Gradient Method

The first adaptive process considered is a gradient based method. This method requires the gradient to be calculable, which inherently requires that there is information available that could directly yield the minimum, the aim of this example is not to demonstrate a viable system - only the concept of gradient based adaption. More advanced methods use methods of gradient estimation that do not require detailed knowledge of the system in question.

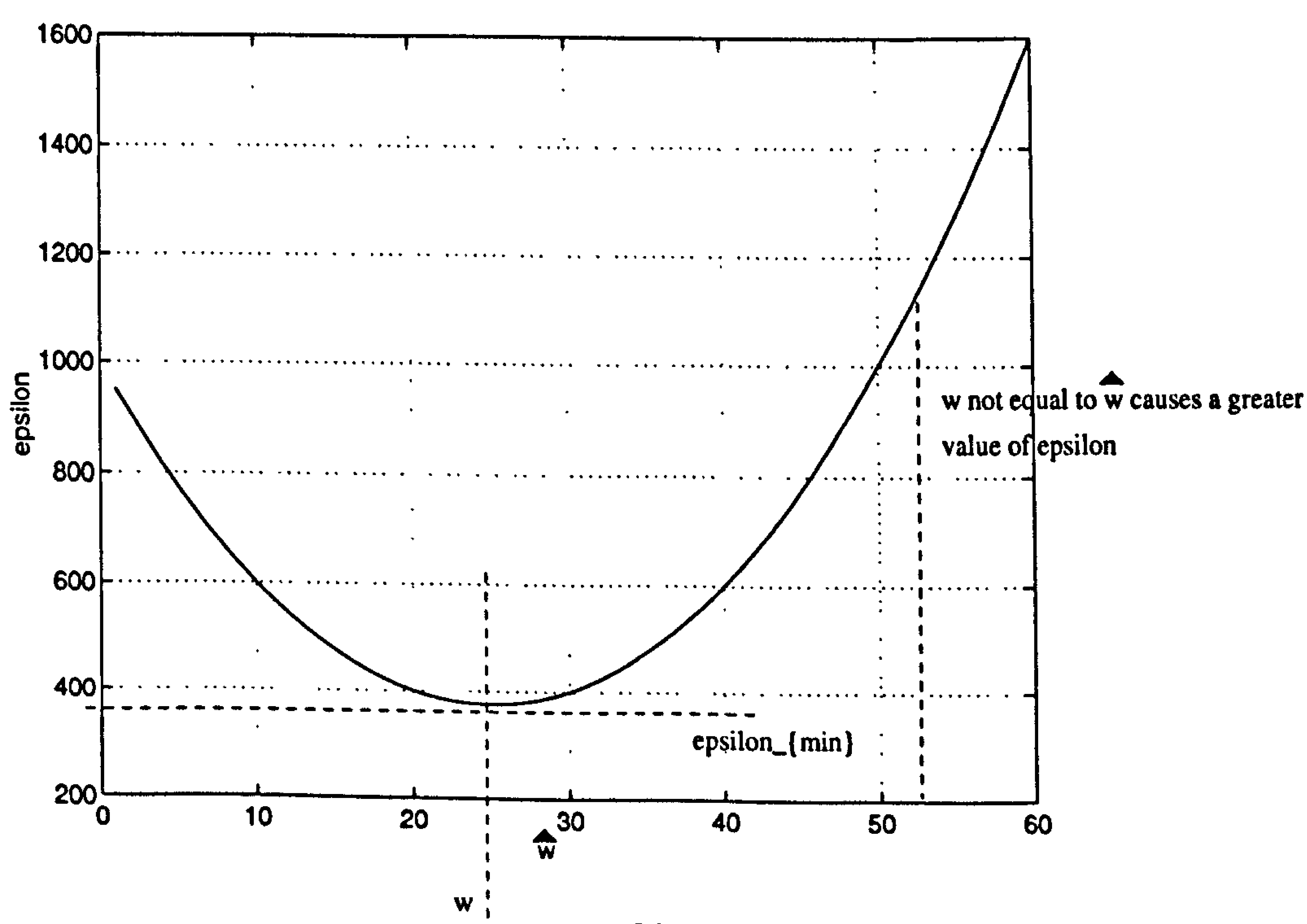
For this example since the system to be considered will have a quadratic performance surface (represented in figure 2.3), the equation for this surface will be of the form:

$$\varepsilon = a\hat{w}^2 + b\hat{w} + c \tag{2.1}$$

This can also be expressed in the form:

$$\varepsilon = \varepsilon_{min} + \lambda\,(\hat{w} - w)^2 \tag{2.2}$$

Figure 2.3: Gradient Search of a Simple Performance Surface



The gradient of this function is:

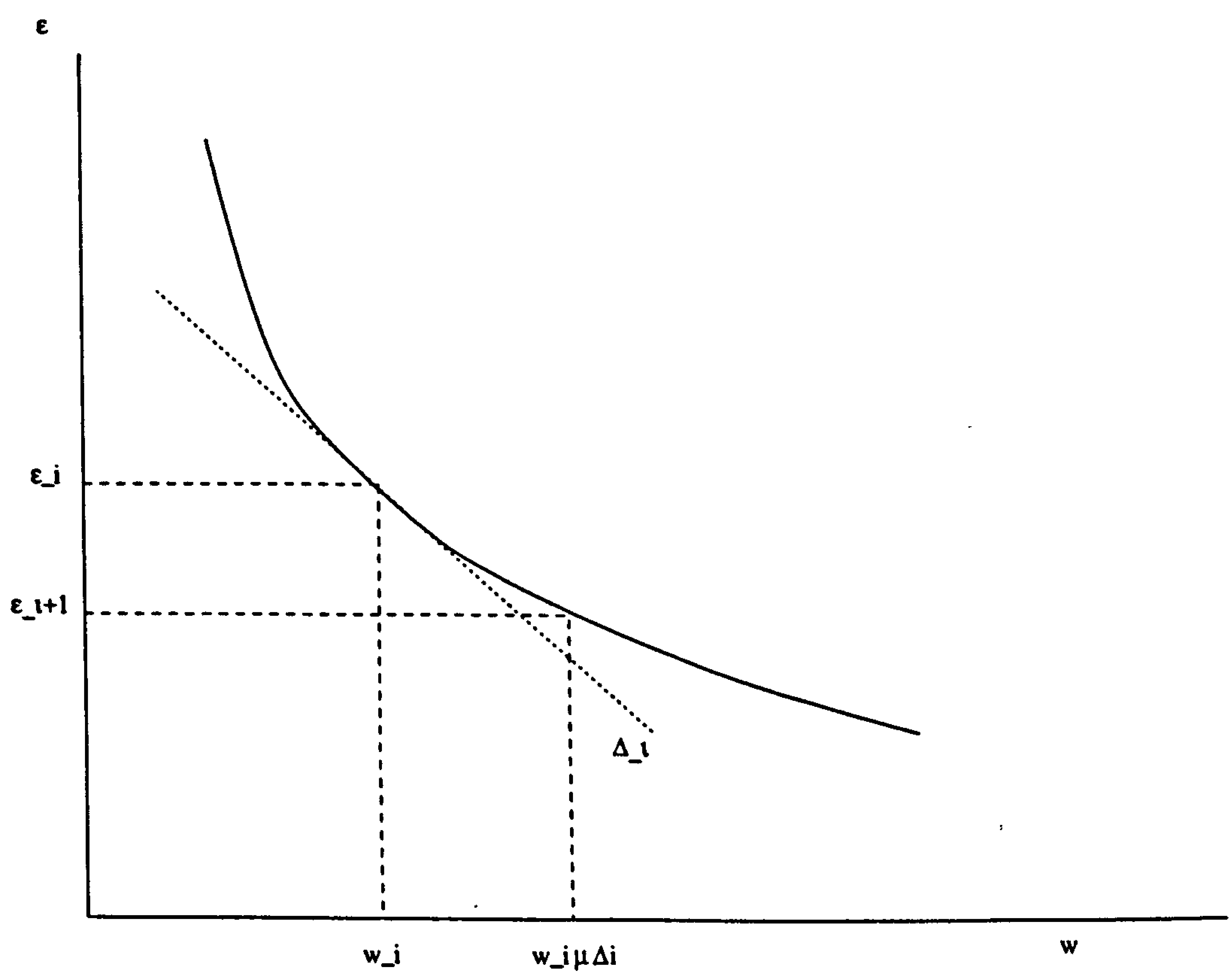
$$\frac{d\varepsilon}{d\hat{w}} = \Delta = 2\lambda\left(\hat{w} - w\right) \tag{2.3}$$

The second derivative, which is constant, is given by:

$$\frac{d^2\varepsilon}{d\hat{w}^2} = 2\lambda \tag{2.4}$$

One method of finding the minimum of the error surface is to start at an arbitrary point, $\hat{w}_0$, and measure the gradient of the surface at this point. A better estimation of the location of the minimum may now be calculated by making $\hat{w}_1$ equal to $\hat{w}_0$ plus an increment proportional to the negative of the gradient. This loop is repeated until a sufficiently good approximation of w is obtained. This process is shown graphically in figure 2.4.

Figure 2.4: Gradient Adaptive Process

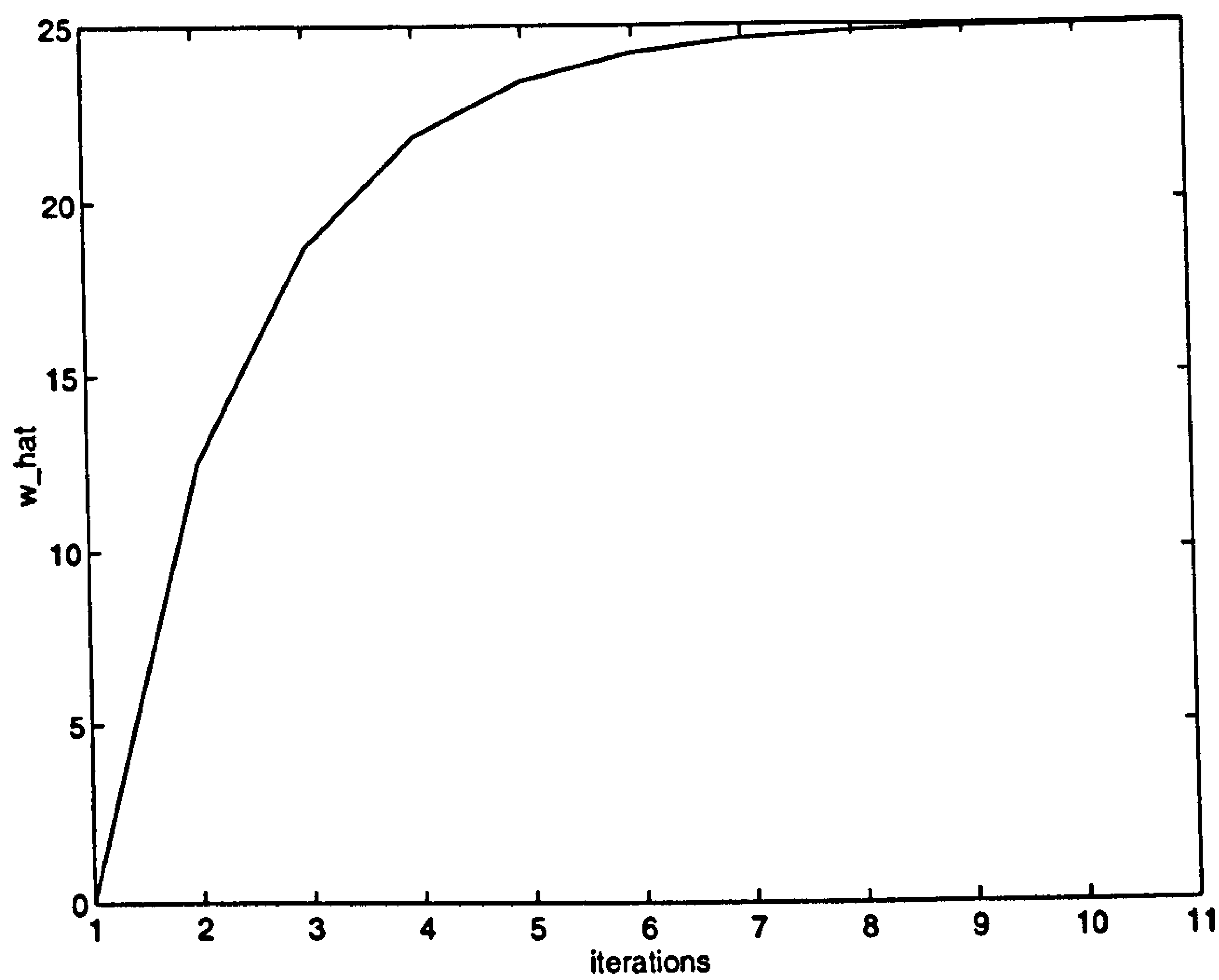


The adaptive process may also be expressed analytically:

$$\begin{aligned}\hat{w}_{i+1} &= \hat{w}_i - \mu\Delta_i && (2.5)\\ &= \hat{w}_i - 2\mu\lambda\left(\hat{w} - w\right) && (2.6)\end{aligned}$$

Given $\mu = 0.25$ the adaption proceeds as below, where the minimum of the function in this case is found at $\hat{w} = 25$.

Figure 2.5: Gradient Adaptive Process in Action



If the value of μ is too high then the adaption process will first become oscillatory and then unstable; the stability criterion can be examined as follows:

$$\hat{w}_1 = \hat{w}_0(1-2\mu\lambda)+2\mu\lambda w \tag{2.7}$$

$$\hat{w}_2 = \hat{w}_1(1-2\mu\lambda)^2+2\mu\lambda w[(1-2\mu\lambda)+1] \tag{2.8}$$

$$\hat{w}_3 = \hat{w}_2(1-2\mu\lambda)^3+2\mu\lambda[(1-2\mu\lambda)^2+(1-2\mu\lambda)+1] \tag{2.9}$$

$$\hat{w}_i = \hat{w}_0(1-2\mu\lambda)^i+2\mu\lambda w\sum_{n=0}^{i-1}(1-2\mu\lambda)^n \tag{2.10}$$

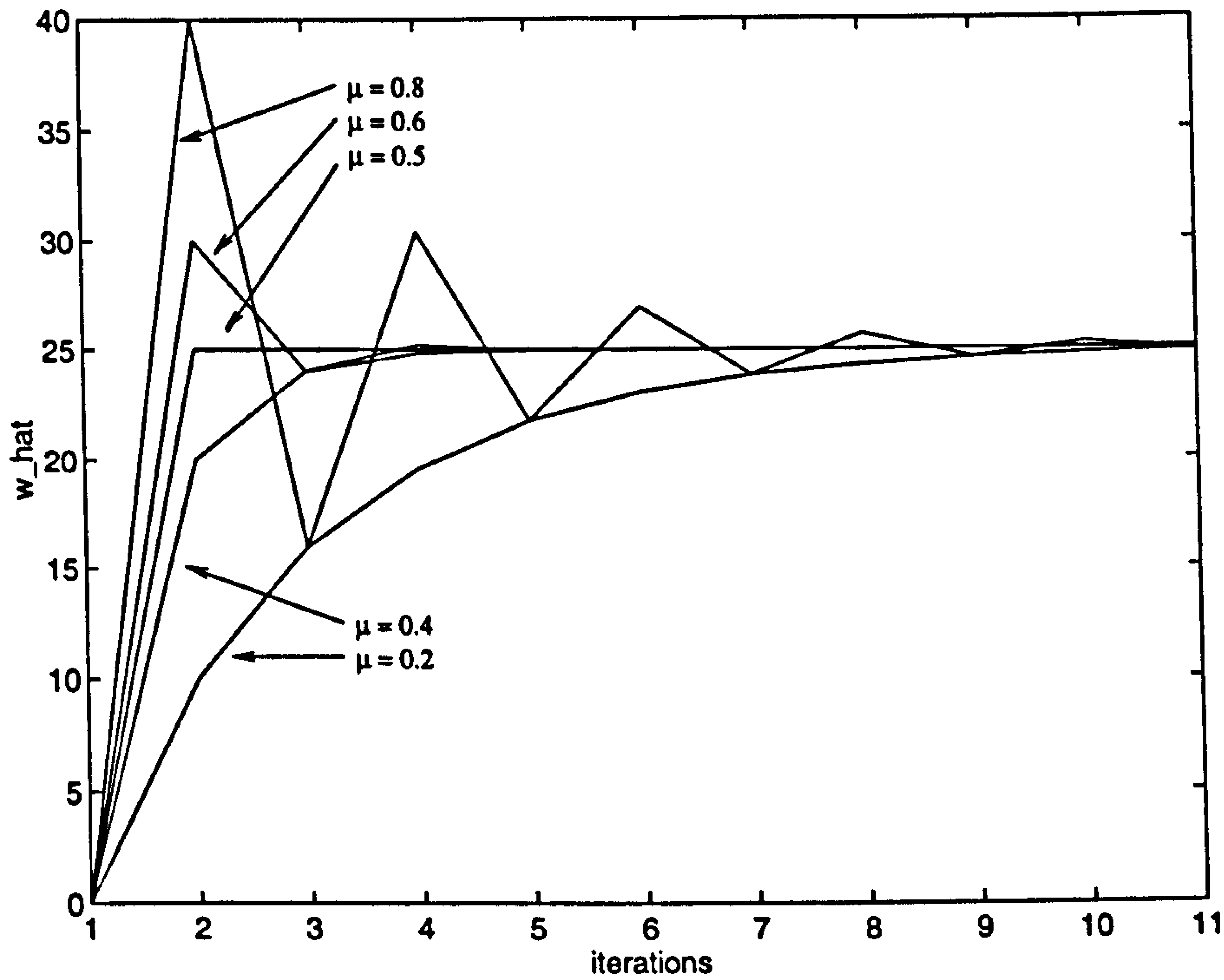
$$= \hat{w}_0(1-2\mu\lambda)^i+2\mu\lambda w\frac{1-(1-2\mu\lambda)^i}{1-(1-2\mu\lambda)} \tag{2.11}$$

$$= w+(1-2\mu\lambda)^i(\hat{w}_0-w) \tag{2.12}$$

It is straight forward to see that this series of values of $\hat{w}$ will only converge if $|(1-2\mu\lambda)|$ is less than unity (Leibnitz' test [9]) and thus the process is only stable within this bound. The quantity $(1-2\mu\lambda)$ is called the geometric ratio.

The effects of the value of μ are interesting; the process uses an error measure that may be negative, thus it is possible for the solution to overshoot, as is demonstrated in the following figure. Values of μ in order of decreasing initial gradient of the solution set are $\mu = 0.8,\ 0.6,\ 0.5,\ 0.4,\ 0.2$, see figure 2.6. Note that the lines on the graph have no true meaning since this is a discrete system, $\hat{w}_i$ only exists on integer values of iteration.

Figure 2.6: Variation in Adaption Rates with μ



When $\mu = 0.5$ the optimum solution is obtained immediately since $(1 - 2\mu\lambda) = 0$ and thus the next approximation is defined by a constant. This special case is called Newton's method.

Returning to equation:

$$\varepsilon_i = \varepsilon_{min} + \lambda(\hat{w}_i - w)^2 \tag{2.13}$$

Substituting for $\hat{w}_i$ yields:

$$\varepsilon_i = \varepsilon_{min} + \lambda(\hat{w}_0 - w)^2(1 - 2\mu\lambda)^{2i} \tag{2.14}$$

Thus ε_i converges towards εmin according to a geometric progression with ratio $(1-2\mu\lambda)^2$. This ratio cannot be negative and thus convergence will not be oscillatory. Convergence is assured by the same constraint as for $\hat{w}$.

2.3.2 Newton's Method

With a gradient search process for a single variable, if the convergence is critically damped and converges in one step with quadratic mean squared error functions then it is called

Newton's method. This name is given due to the similarity that exists between the problem described here and the method of finding roots of a polynomial known as Newton's method.

Newton's method allows the approximate solution to problems of the nature of $f(x) = 0$, yielding the exact solution is a single iteration for a quadratic polynomial. The method consists of starting with an initial guess, x_0, and then using the first derivative, $f'(x_0)$, to compute the next estimate x_1. x_1 is found where the tangent to the polynomial at x_0 intersects with the x-axis. In general then:

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \tag{2.15}$$

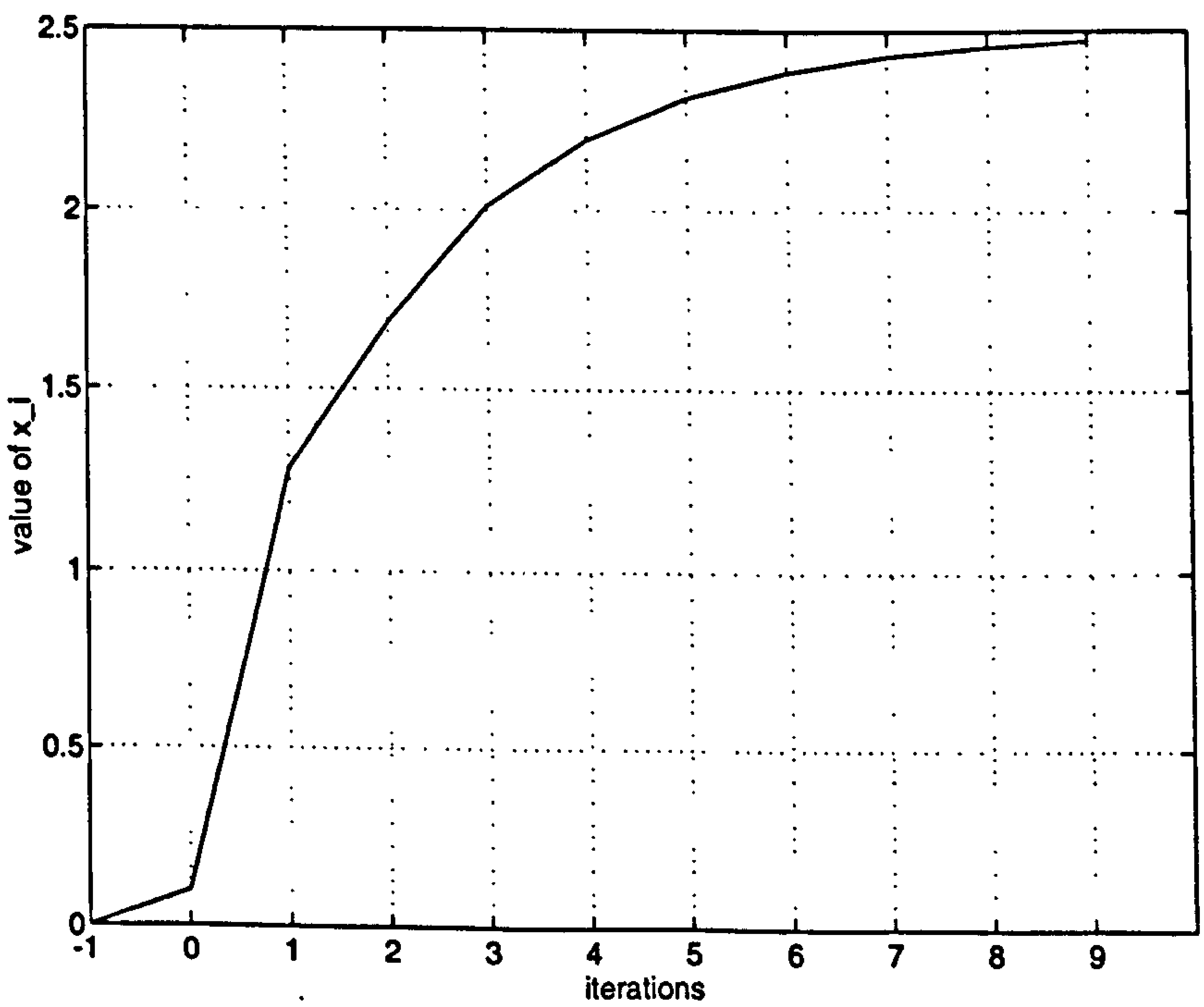
The convergence of Newton's method depends on the initial guess and on the function f. For a large class of functions the method is known to converge quickly but there are functions where, given a poor initial guess, the method will not converge.

So far the method assumed that the function was continuous and that the derivative was explicitly available as a function of the variable, it is possible to construct a discrete version that uses a gradient estimate based of the values available:

$$x_{k+1} = x_k - \frac{f(x_k)(x_k - x_{k-1})}{f(x_k) - f(x_{k-1})} \tag{2.16}$$

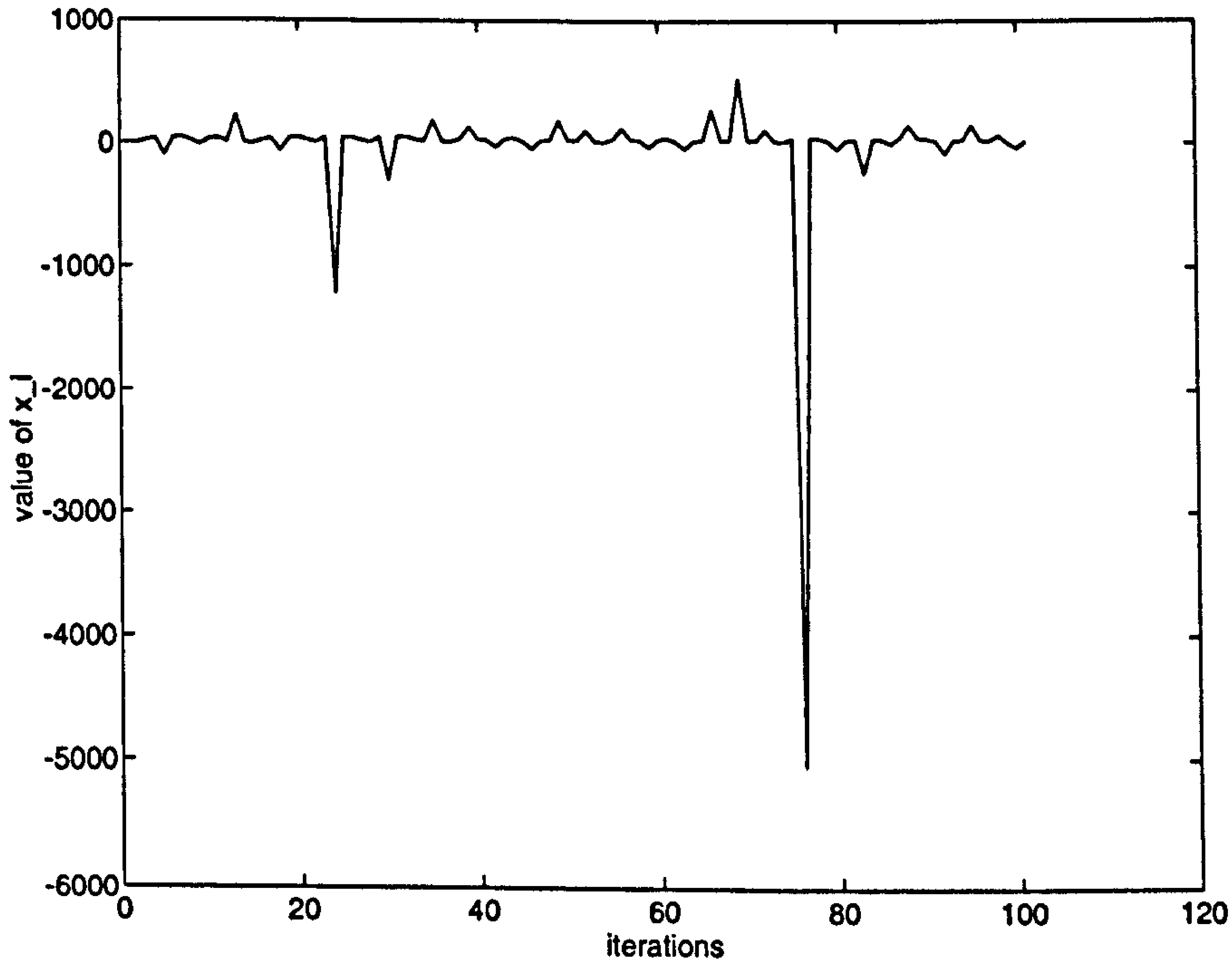
As a demonstration the minimum of the quadratic $f(x) = 0.16x^2 - 0.8x + 1$ is found using $x_{-1} = 0$ and $x_0 = 1$ as the initial values and the discrete version of Newton's method, see figure 2.7.

Figure 2.7: Newton's Method



The frailty of the method can be demonstrated by attempting to find the minimum of the quadratic $f(x) = x^2 - 50x + 1000$ with the same initial points. This problem has already been solved using the exact gradient. The figure 2.8 demonstrates the process is sporadically deviating further from the correct solution as the number of iterations completed increases. This is a problem encountered when using Newton's method because the gradient used is an estimate and the process attempts to arrive at the solution in a single iteration. In practice methods that converge more slowly are more desirable since they are more robust to gradient estimation noise.

Figure 2.8: The Frailty of Newton's Method



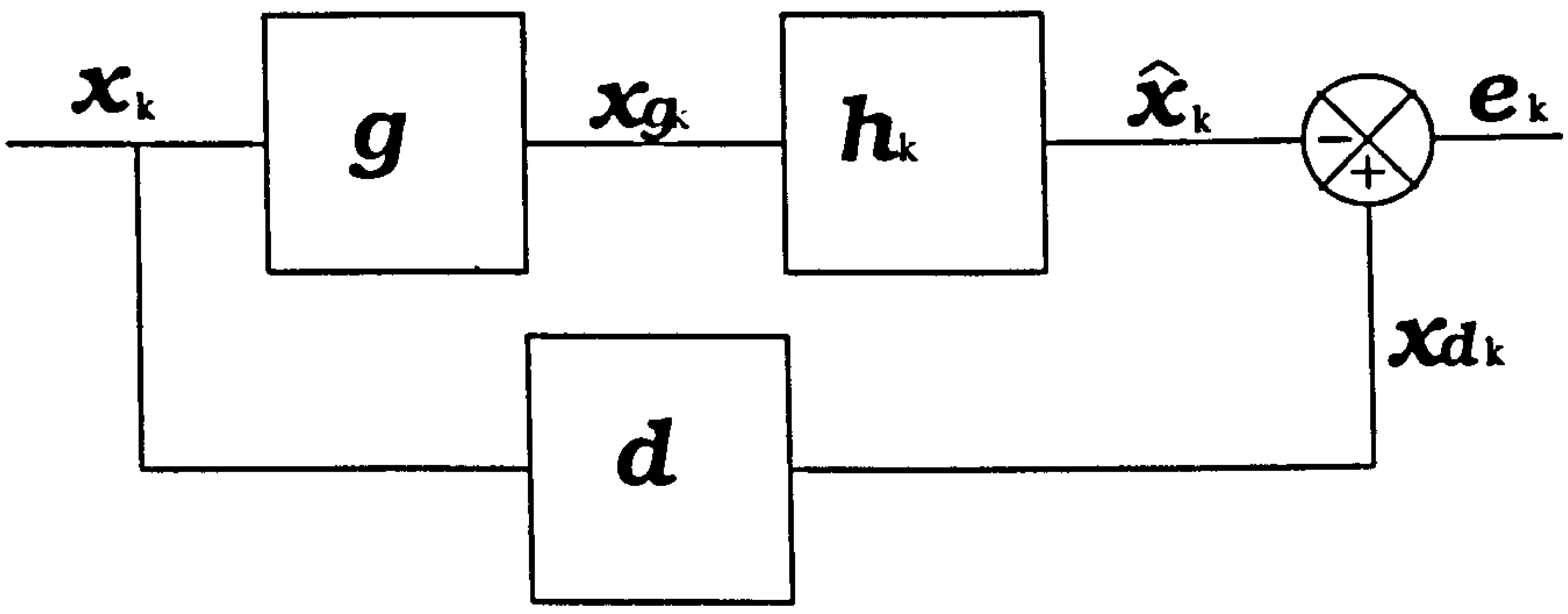
2.4 Adaptive Inverse Filtering

Inverse filtering is important to room acoustic correction and thus it is used as an example to develop the Least Mean Squares (LMS) algorithm. A block diagram of an adaptive inverse filter in given in figure 2.9. In this diagram g is a FIR approximation to the impulse response of a room and $\mathbf{h}_k$ is the k^{th} iteration of the inverse filter. A FIR model of the desired system response is set in $\mathbf{d}$. $\mathbf{x}_k$ is the signal value at the k^{th} iteration etc. and $\mathbf{e}_k$ is the error, which is the performance measure used to driver the adaption.

The aim on an adaptive inverse filter is to adaptively attempt develop a filter to completely reverse the effects of a process.

2.4.1 Justification for the NLMS Algorithm

Figure 2.9: Adaptive Inverse Filter Block Diagram



The filters g, h and d are FIR filters that are represented by the vectors. (Note that the subscript has been dropped from $\mathbf{h}_k$ since it is initially being considered as invariant). It is important to note that in some systems g may not be of finite length, for instance where g represents a room acoustic transfer function.

If g were to be truncated at some point then it is possible to represent the problem in terms of matrices.

$$\mathbf{d} = \begin{bmatrix} d_0 \\ d_1 \\ \vdots \\ d_{L_g+L_h+1} \end{bmatrix} \qquad \mathbf{g} = \begin{bmatrix} g_0 \\ g_1 \\ \vdots \\ g_{L_g} \end{bmatrix} \qquad \mathbf{h} = \begin{bmatrix} h_0 \\ h_1 \\ \vdots \\ h_{L_h} \end{bmatrix} \tag{2.17}$$

In order to allow for the convolution of g and h, so that a filter to represent their combined effects is produced, the matrix G is defined:

$$
\mathbf{G} = \begin{bmatrix} g_0 & & & \\ g_1 & g_0 & & 0 \\ \vdots & g_1 & \ddots & \\ g_{L_g} & \vdots & & g_0 \\ & g_{L_g} & & g_1 \\ 0 & & \ddots & \vdots \\ & & & g_{L_g} \end{bmatrix} \tag{2.18}
$$

Thus ideally:

$$
\mathbf{d} = \mathbf{G}\mathbf{h} \tag{2.19}
$$

Due to the nature of $\mathbf{G}$ there will be no exact solution to this equation. (If this equation is compared with the matrix expression of a set of simultaneous equations it can be seen that there are more equations than there are variables and hence generally there will be no exact solution.) The optimum value of $\mathbf{h}$ can be calculated using:

$$
\mathbf{h}_{opt} = \left[\mathbf{G}^{\mathbf{T}}\mathbf{G}\right]^{-1}\mathbf{G}^{\mathbf{T}}\mathbf{d} \tag{2.20}
$$

This will be referred to as the matrix method. The derivation of this result can be found in Appendix 9.3.

2.4.2 Adaptive Inverse Filter Without Truncating "g"

If g cannot be expressed in truncated form then the only alternative is to assume that there will be knowledge of x_{gk}, which is in fact the case for inverse room acoustic filtering. If a vector $\mathbf{x}_{gk}$ is constructed:

$$
\mathbf{x}_{gk} = \begin{matrix} x_{gk} \\ x_{gk-1} \\ \vdots \\ x_{gk-L_h} \end{matrix} \tag{2.21}
$$

Then $\hat{x}_k$ is now given by:

$$
\hat{x}_k = \mathbf{x}_{gk}^T\mathbf{h} = (\mathbf{h})^T\mathbf{x}_{gk} \tag{2.22}
$$

And thus the error can be calculated.

$$
e_k = x_{dk} - \mathbf{x}_{gk}^T\mathbf{h} \tag{2.23}
$$

Squaring to get the instantaneous squared error:

$$e_k^2 = x_{dk}^2 + (\mathbf{h})^T \mathbf{x}_{gk} \mathbf{x}_{gk}^T \mathbf{h} - 2x_{dk}(\mathbf{h})^T \mathbf{x}_{gk} \tag{2.24}$$

Assume that e_k, x_{dk} and $\mathbf{x}_{gk}$ are statistically stationary and take expected value over k:

$$E[e_k^2] = E[x_{dk}^2] + (\mathbf{h})^T E[\mathbf{x}_{gk}\mathbf{x}_{gk}^T]\mathbf{h} - 2E[x_{dk}\mathbf{x}_{gk}^T]\mathbf{h} \tag{2.25}$$

Note that while the expected value of any sum is the sum of the expected values the expected value of a product is only the product of the expected values when the variables are independent. The variables $\mathbf{x}_k$ and x_{dk} are not independent in this case. Defining $\mathbf{X}_{gk}$ as:

$$\mathbf{X}_{gk} = E[\mathbf{x}_{gk}\mathbf{x}_{gk}^T] \tag{2.26}$$

Thus $\mathbf{X}_{gk}$ is a square matrix typically known as the input correlation matrix. Also define $\mathbf{P}_{gk}$:

$$\mathbf{P}_{gk} = E[x_{dk}\mathbf{x}_{gk}] \tag{2.27}$$

Returning to the expression for the MSE, but substituting $\mathbf{X}_{gk}$ and $\mathbf{P}_{gk}$ as appropriate:

$$E[e_k^2] = E[xdk^2] + (\mathbf{h})^T \mathbf{X}_{gk}\mathbf{h} - 2\mathbf{P}_{gk}^T\mathbf{h} \tag{2.28}$$

The mean squared error is in the form of a quadratic error surface when the inputs are statistically stationary.

If this expression for the MSE is differentiated then the gradient of the error surface is produced. The differentiation is is not considered in detail here but is of a very similar form to that used in the derivation of the matrix method, Appendix 9.3.

$$\Delta = \frac{\partial E[e_k^2]}{\partial \mathbf{h}} \tag{2.29}$$

$$= 2\mathbf{X}_{gk}\mathbf{h} - 2\mathbf{P}_{gk} \tag{2.30}$$

This expression of the gradient is difficult to work with. Practical algorithms use an estimate of this error. Much of the performance of an algorithm depends on the qualities of this gradient estimate.

If we now take a steepest descent algorithm, which is of the form:

$$\mathbf{Gh}_{k+1} = \mathbf{Gh}_k - \mu\hat{\Delta}_k \tag{2.31}$$

And for an estimate of the gradient of the performance surface, take e_{dk}^2 instead of $E[e_{dk}^2]$. With this approximation with the partial differentiation carried out as before:

$$\hat{\Delta} = \frac{\partial e_k^2}{\partial \mathbf{h}} \tag{2.32}$$

$$= 2\mathbf{x}_{gk}\mathbf{x}_{gk}^T\mathbf{h} - 2x_{dk}\mathbf{x}_{gk} \tag{2.33}$$

Since $\mathbf{x}_{gk}^T\mathbf{Gh}$ is a scalar it is permissible to move this in front of the preceding $\mathbf{x}_k$ term. Note that the multiplication is no longer fully conformable, the multiplication of $\mathbf{x}^T\mathbf{Gh}$ must occur first, this is implicit in the left to right ordering.

$$\hat{\Delta} = 2(\mathbf{x}_{gk}^T\mathbf{h})\mathbf{x}_{gk} - 2x_{dk}\mathbf{x}_{gk} \tag{2.34}$$

$$= -2(x_{dk} - \mathbf{x}_{gk}^T\mathbf{h})\mathbf{x}_{gk} \tag{2.35}$$

Now substituting e_{dk} for $x_{dk} - \mathbf{x}_{gk}^T\mathbf{h}$ gives:

$$\hat{\Delta} = -2e_{dk}\mathbf{x}_{gk} \tag{2.36}$$

Thus the tap update equation becomes:

$$\mathbf{h}_{k+1} = \mathbf{h}_k + 2\mu e_k \mathbf{x}_{gk} \tag{2.37}$$

This expression can be applied directly to inverse filter type problems such as room acoustic equalization etc. There is one additional modification that is normally carried out. In a real system the value of the signal may at times fall well below the peak and when it does the adaption rate falls with it; but the value of μ cannot be increased since if a large signal were to occur then the system would potentially become unstable. (There is a limited range of values that μ can take for assured stability.) The problem is compounded by the fact that the value of the error is largely dependent of the value of $\mathbf{x}_{gk}$ as the filter begins to converge.

The solution is simply to divide the vector $\mathbf{x}_{gk}$ by its magnitude squared, to take into account the two different effects of reduced levels of x_{gk}, ie. to normalize it. It is also normal to replace 2μ with β. A constant K is included to prevent the normalization becoming extreme in the event of signal levels falling to very low levels, this is important since otherwise the effectively elevated noise levels would damage the filter. The concept is not surprisingly called the Normalized Least Mean Squares (NLMS) algorithm.

$$\mathbf{h}_{k+1} = \mathbf{h}_k + \beta \frac{e_k \mathbf{x}_{gk}}{\mathbf{x}_{gk}^T \mathbf{x}_{gk} + K} \tag{2.38}$$

2.4.3 Properties of the NLMS Algorithm

In practice the NLMS algorithm has been found to be stable, reliable and produce a good estimate of the filter required in a large number of different situations. With the normalizing addition there is no need to consider the signal levels, the algorithm will adapt well regardless.

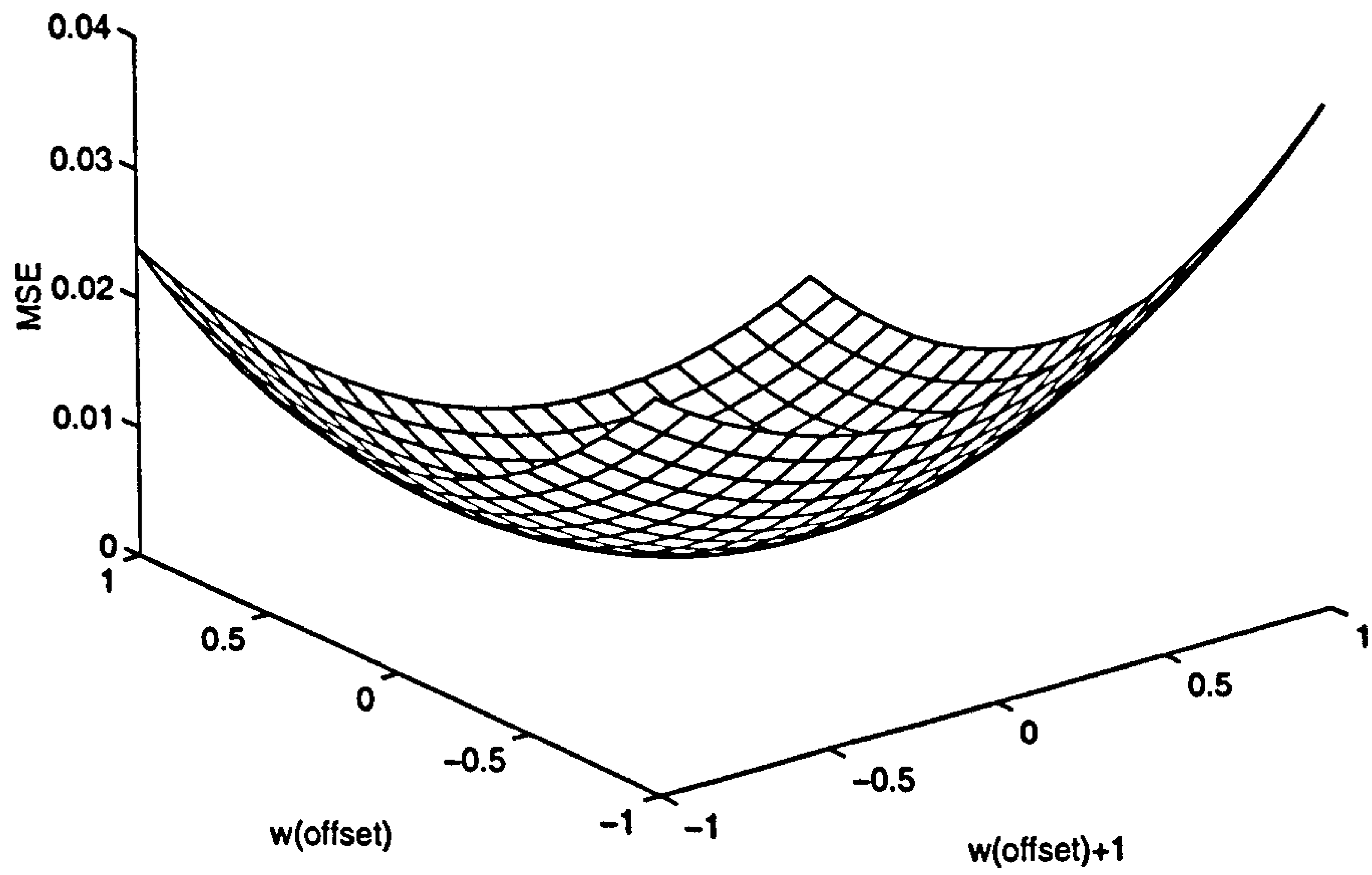
In general values of β around 0.6 give good results, though the value can be raised normally to about 1.2 before instability results. It is possible to demonstrate analytically the range of β over which the algorithm is stable but the derivation is not significant to the remainder of this work and therefore the reader is referred to [62].

2.4.4 The NLMS Algorithm in Action

In order to demonstrate the algorithm in action a truncated filter g was used. (The filter in fact is a truncated and limited bandwidth room impulse response but that is of no significance here.) Of the large number of weights in the test filter two were selected at random to be displayed.

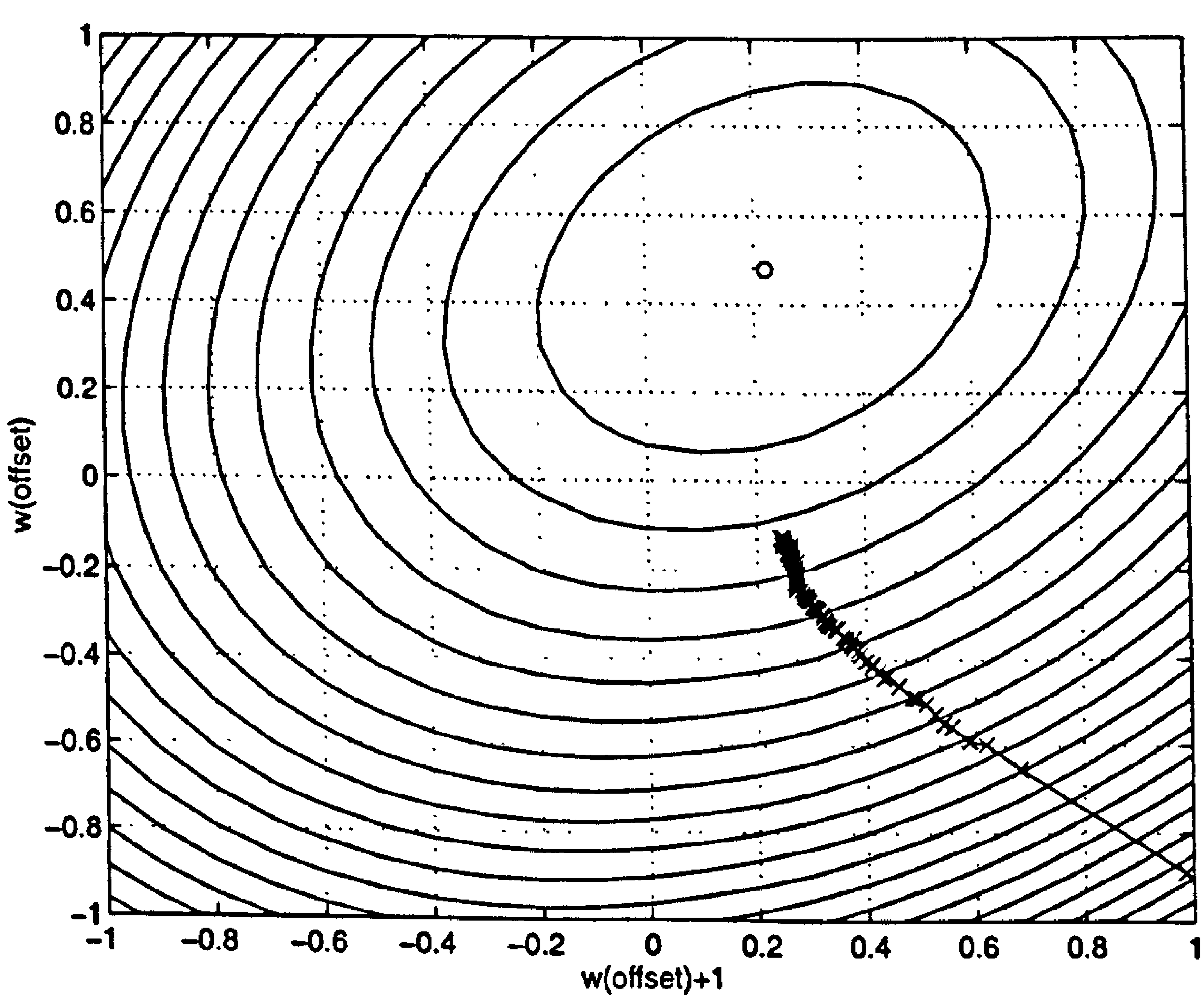
First the optimum weight vector was calculated using the solution derived in the section titled "Justification for the Matrix Method." The the mean squared error was calculated for the entire filter when the chosen two weights were perturbed over an area. The results demonstrate the error surface and may be seen in figure 2.10

Figure 2.10: Variation of MSE with the Weight Vector



If the NLMS algorithm is now used to find a solution for the same problem and the changing values of the two elements of the weight vector are recorded at each length(filter) iterations then the maturation of the filter can be observed. The two filter weights at each stage in the development of the filter are plotted on top of a set of contours of the error surface to allow easy comparison. The convergence is initially rapid but slows as the filter approaches the ideal solution. The plot is given in figure 2.11.

Figure 2.11: Successive Generations of the Estimated Filter



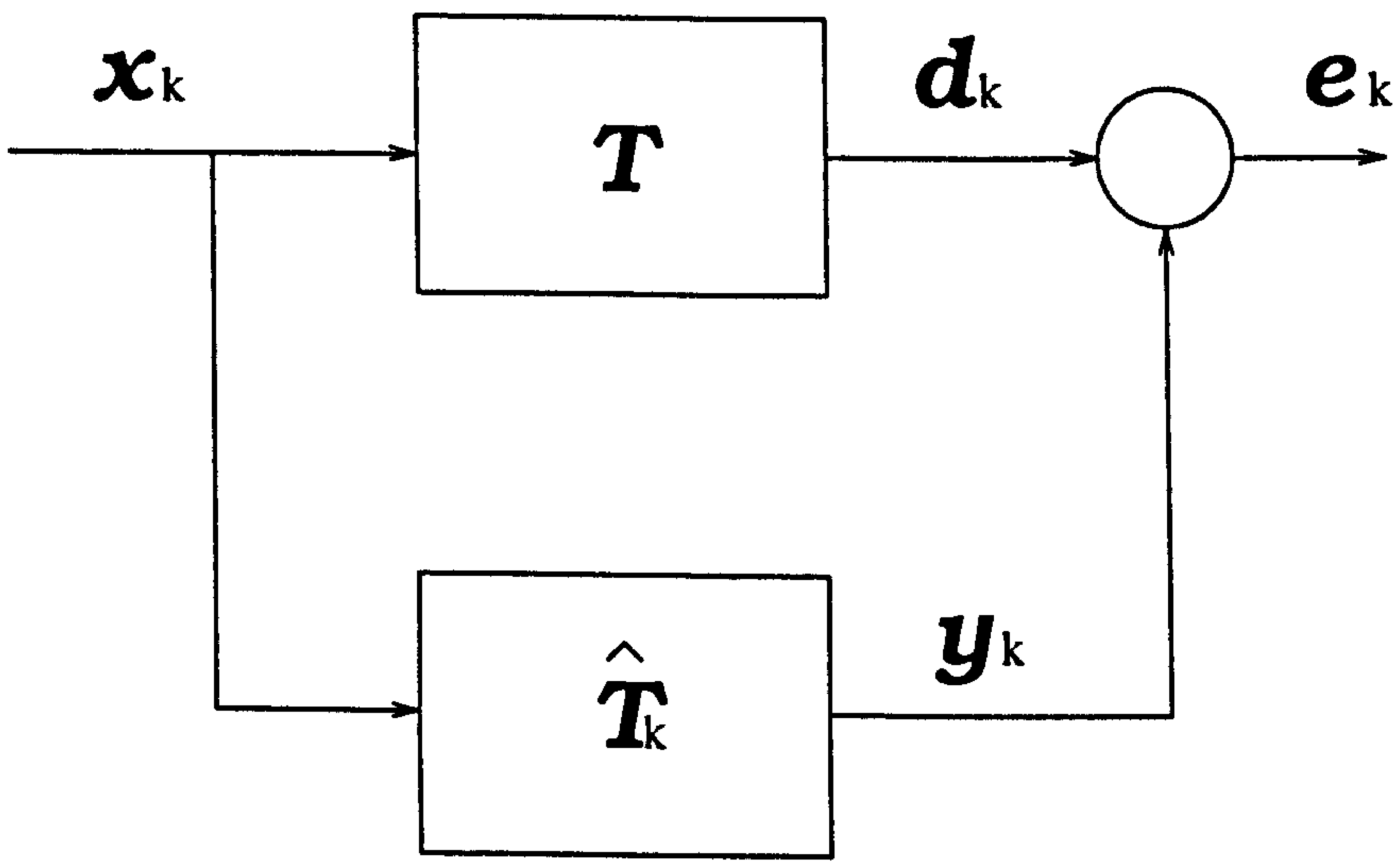
2.5 Applying the NLMS Algorithm to System Identification

The system identification problem has received a great deal of attention since it is important to many commercial applications, for instance echo cancellation for video conferencing depends on it. Normally the form of the system model is chosen to be an FIR filter since this makes the problem more tractable and stability easier to assure. However, for the control problem developed in a later chapter, it will be necessary to to look at the slightly more complicated case where the model form is both an FIR and an IIR filter. The derivation of a solution below follows Widrow's analysis [62].

A block diagram for a typical system identification setup is given in figure 2.12.

$\mathbf{T}$ is the system to be identified and thus $\mathbf{d}_k$ is the desired output of the system for an input of $\mathbf{x}_k$. $\hat{\mathbf{T}}_k$ is the model of the system at iteration k, which produces output $\mathbf{y}_k$. The difference between $\mathbf{d}_k$ and $\mathbf{y}_k$ is the error, which is the driving force of the process.

Figure 2.12: System Identification Block Diagram



Starting with a transfer function of the form:

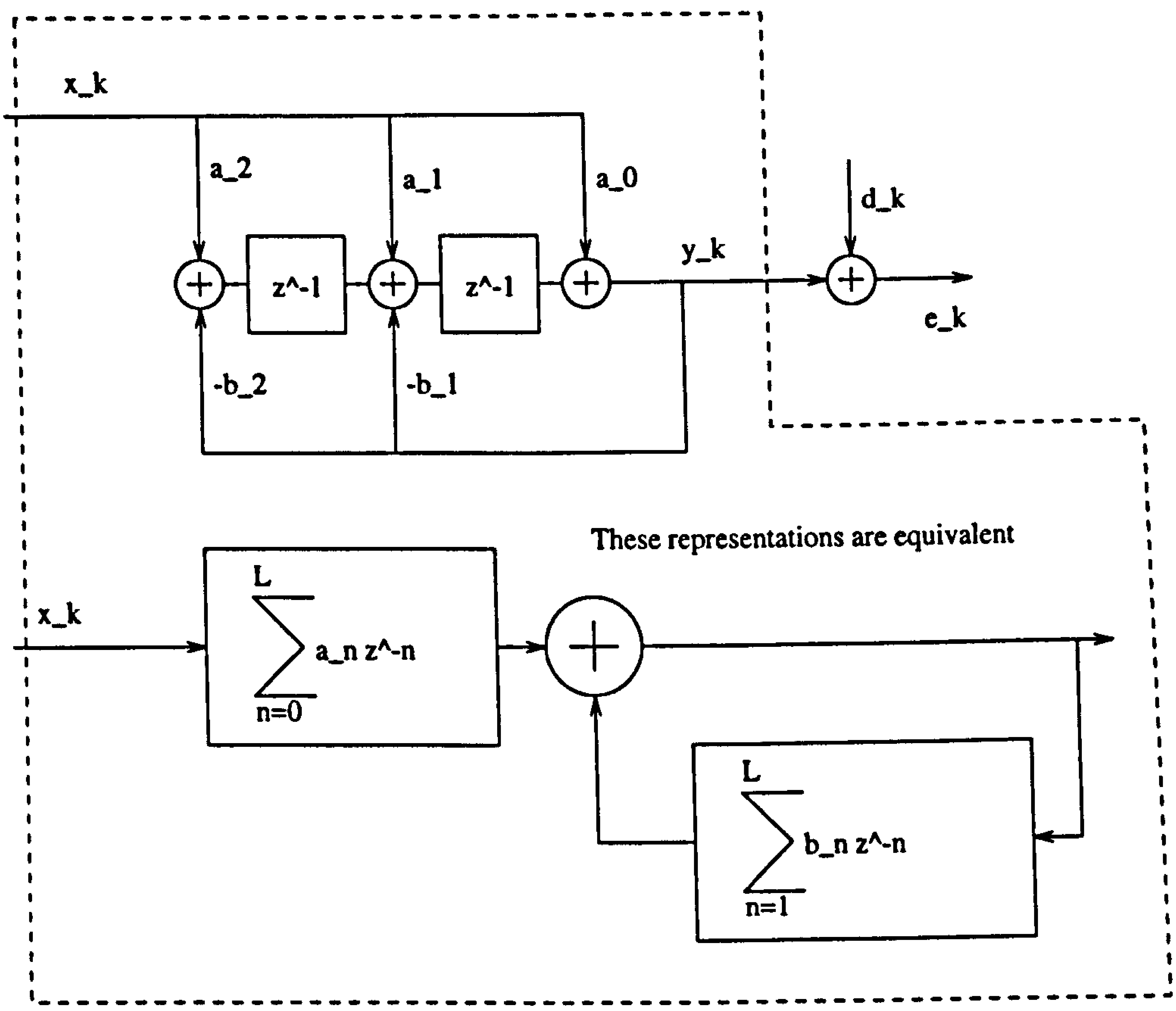
$$\hat{T}(z) = \frac{a_{nk}z^{-n} + a_{n-1k}z^{k-1} + \cdots + a_0}{b_{nk}z^{-k} + b_{n-1k}z^{k-1} + \cdots + 1} \tag{2.39}$$

Where a_{nk} is the n^{th} tap of $\mathbf{a}_k$ at sample number k and a_{n-1k} is the $n-1^{th}$ tap of $\mathbf{a}_k$ etc. This may also be expressed as the function below for an input x_k and an output y_k.

$$y_k = \sum_{n=0}^{L} a_n k x_{k-n} + \sum_{n=1}^{L} b_n k y_{k-n} \tag{2.40}$$

This can also be expressed diagrammatically, see figure 2.13:

Figure 2.13: System Identification Diagrams



Defining:

$$\mathbf{w}_k = [a_0k \cdots a_Lk b_0k \cdots b_Lk]^T \tag{2.41}$$

$$\mathbf{u}_k = [x_k \cdots x_{k-L} y_k \cdots y_{k-L}]^T \tag{2.42}$$

From knowledge of the system the following can be written:

$$e_k = d_k - y_k \tag{2.43}$$

$$= d_k - \mathbf{w}_k^T \mathbf{u}_k \tag{2.44}$$

Considering $\frac{\partial e_k^2}{\partial \mathbf{w}_k}$ as an approximation to the performance surface gradient in order to use a steepest descent type algorithm.

$$\hat{\Delta}_k = \frac{\partial e^2}{\partial \mathbf{w}_k} = 2e\frac{\partial e}{\partial \mathbf{w}_k} \tag{2.45}$$

$$= 2e_k \left[\frac{\partial e_k}{\partial a_0 k} \cdots \frac{\partial e_k}{\partial a_L k}\frac{\partial e_k}{\partial b_1 k} \cdots \frac{\partial e_k}{\partial b_L k}\right]^T \tag{2.46}$$

$$= -2e_k \left[\frac{\partial y_k}{\partial a_0 k} \cdots \frac{\partial y_k}{\partial a_L k}\frac{\partial y_k}{\partial b_1 k} \cdots \frac{\partial y_k}{\partial b_L k}\right]^T \tag{2.47}$$

Defining:

$$\alpha_{nk} \triangleq \frac{\partial y_k}{\partial a_n} \tag{2.48}$$

$$\beta_{nk} \triangleq \frac{\partial y_k}{\partial b_n} \tag{2.49}$$

Using the definition of y_k

$$\alpha_{nk} = x_{k-n} + \sum_{l=1}^{L} b_l \frac{\partial y_{k-1}}{\partial a_n} \tag{2.50}$$

$$= x_{k_n} + \sum_{l=1}^{L} b_l \alpha_{nk-l} \tag{2.51}$$

$$\beta_{nk} = y_{k-n} + \sum_{l-1}^{L} b_l \frac{\partial y_{k-l}}{\partial b_n} \tag{2.52}$$

$$= y_{k-n} + \sum_{l=1}^{L} b_l \beta_{nk-l} \tag{2.53}$$

Thus $\hat{\Delta}_k$ can now be written as:

$$\hat{\Delta}_k = -2e_k[\alpha_{0k} \cdots \alpha_{Lk} \beta_{1k} \cdots \beta_{Lk}]^T \tag{2.54}$$

Using a steepest decent type of formula the LMS algorithm can therefore be written as:

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \mathbf{C}\hat{\Delta}_k \tag{2.55}$$

The convergence constant μ has been replaced by diagonal matrix C, which allows the convergence factors for each b to be controlled individually. C is defined as follows:

$$C = \begin{bmatrix} \mu & 0 & & \cdots & & 0 \\ 0 & \ddots & & & & \\ & & \mu & & & \vdots \\ \vdots & & & \nu_1 & & \\ & & & & \ddots & 0 \\ 0 & 0 & & \cdots & 0 & \nu_L \end{bmatrix} \tag{2.56}$$

Chapter 3

An Introduction to the Stereo System and Human Hearing

3.1 The Reproduction Problem

The human hearing is a well developed sense which makes the design of any system to produce convincing audio illusion of a complicated source a considerable engineering challenge. Humans can hear from about 2Hz (though 20Hz is required for a sense of tonality) to about 20kHz and over a dynamic range of about 140dB. (0dB is equivalent to a vibration about the size of a hydrogen molecule.) [22]. (There are animals that have hearing capabilities exceeding those of humans in almost every way except perhaps when discriminating subtle changes in frequency and loudness [16].)

The ideal sound reproduction system would reproduce the stimulus sufficiently accurately to be able to completely fool the human hearing sense. Even simply designing a transducer capable of achieving a dynamic range of 140dB and a frequency response of 2Hz - 20kHz represents a significant problem. The reproduction must also be sufficiently accurate for a wide range of concurrent signals to satisfy the human hearing system's ability to detect subtle changes in intensity and tone - which means that the output of the reproducing devices needs to be ideally linear over a wide range of frequencies and free from harmonic distortion. In addition to these basic abilities the human hearing system provides the ability to locate the sound source which will also need to be satisfied if ideal reproduction is to be achieved. The moving coil loudspeaker has long been the standard for audio reproduction, though at present even the better examples are not able to fully satisfy the human hearing ability. Moving coil loudspeakers, however, reproduce sound only at a single effective point. Ideally a large number of loudspeakers should be used to reproduce the locations of the original sound sources; but this is not practical because of economic considerations. Blumlein [54] proposed a system that has since become accepted as the standard for audio reproduction, the "stereo" system, which uses only two loudspeakers for reproduction.

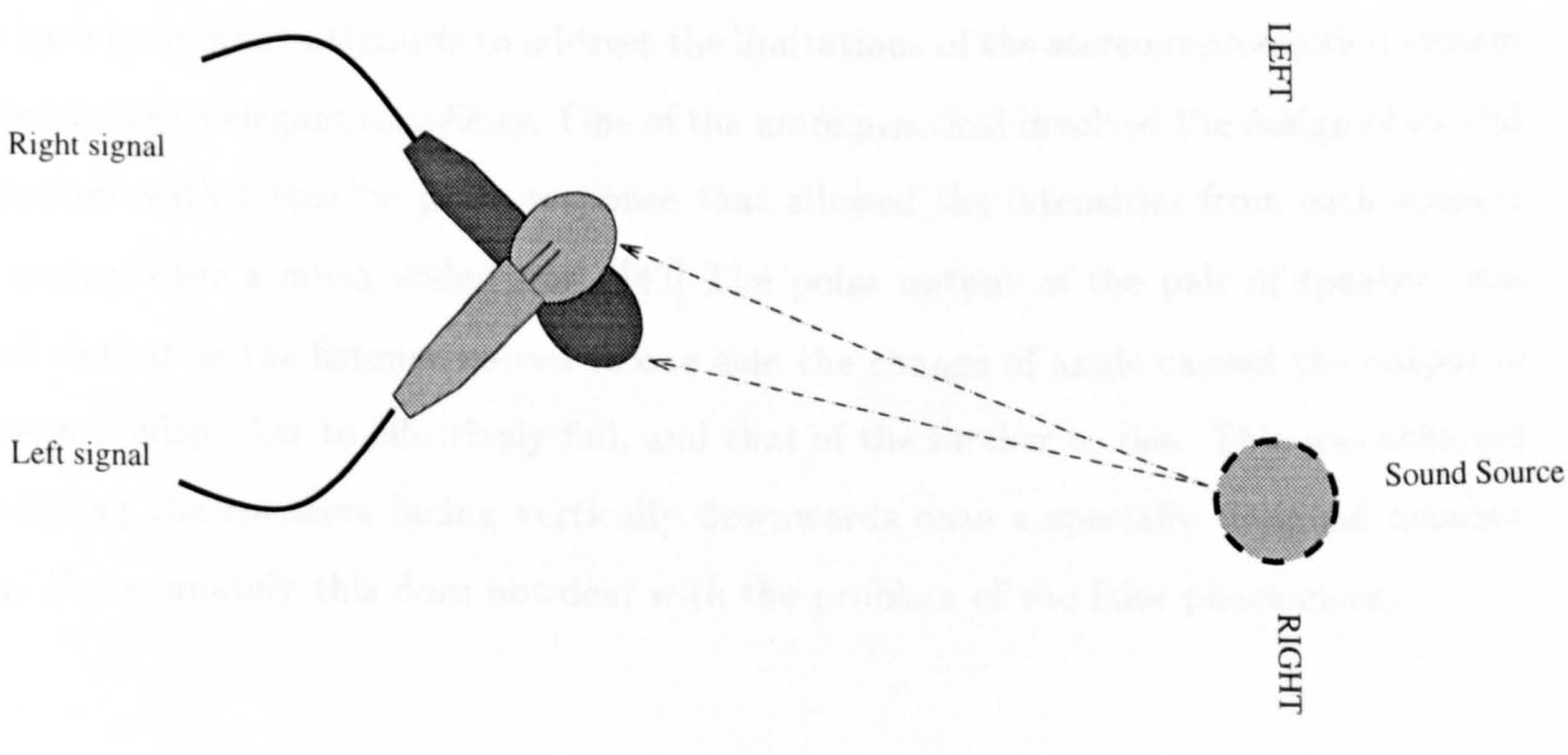
3.2 The "Stereo" Reproduction System

The stereo audio reproduction system that we are familiar with was developed by Alan Blumlein in the early 1930s. Blumlein abandoned all hope of reproducing both the phase and intensity clues at the ears of the listener (both are used for sound location) and instead converted the phase differences to intensity differences using special circuits he devised which he called "shufflers" [54].

Current music recorded for stereo reproduction may use different techniques but the objective remains unchanged - the recorded signals have intensities relating to the position of the signal source in the stereo image. (The stereo system is incapable of reproducing phase clues.) The human hearing system makes use of both phase and intensity clues to locate a sound and thus the stereo system is only capable of a partial illusion. The relative importance that the human hearing system places on the two location methods and the frequency range over which each operate can be used to estimate the effectiveness of the stereo concept and define the requirements that a stereo reproduction system must meet to maintain the stereo illusion.

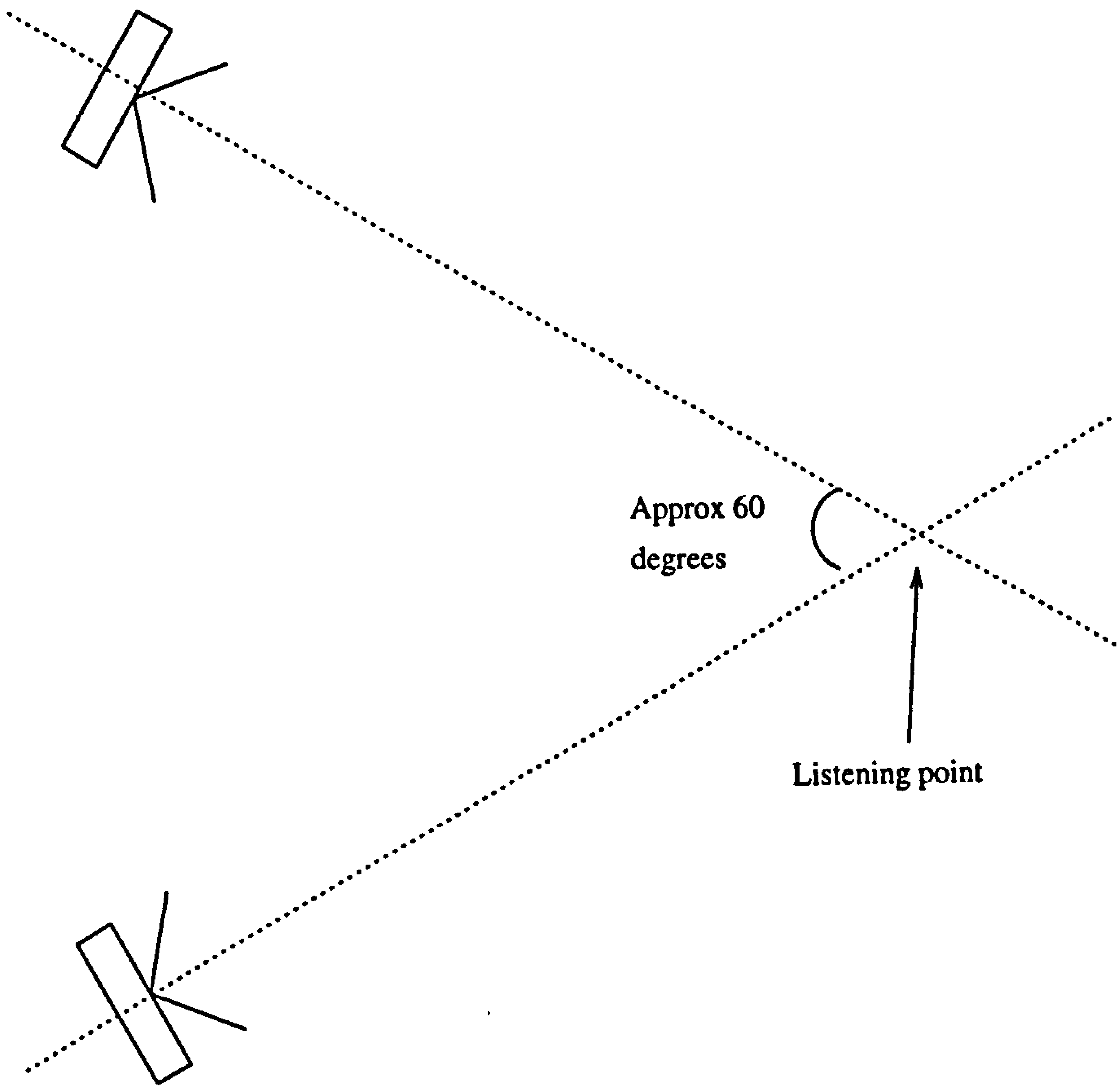
Stereo sound is often currently recorded using crossed cardioid microphones, here the intensity differences are produced by the cardioid microphone's reduced sensitivity off-axis. This method is shown the diagram. (The effect may also be produced using a mid-side arrangement, which consists of an omni-directional microphone facing forwards and a figure 8 microphone facing horizontally perpendicular to this. Due to the nature of figure 8 microphones the microphone is equally sensitive front and back but the result of a signal from the rear is out of phase by 180 degrees when compared to one on the front face. Thus when the signals from the two microphones are added, the result is a pickup pointing to one side, and when they are subtracted it is pointing to the other. There are advantages claimed for this method in terms of the final quality of the recording [35].) In practice many current stereo recordings are not produced by recording the sound at a single point at all, but by recording each instrument individually and constructing a virtual stereo image by means of adjusting each signal's relative intensity on the two channels.

Figure 3.1: A Stereo Recoding Method



Stereo sound is reproduced through two loudspeakers angled at roughly 30^o to each side of the listener and separated by a reasonable distance [54] *pp 34*. (The distance between the speakers has no direct effect on the operation of the system provided that the value is reasonable, typically a value of 3 to 5m is used in a normal system.) There is a single point at which the reproduction will be optimal, deviation from this point to the side brings the listener closer to one of the speakers and upsets the balance of intensity and introduces false phase clues. Deviation towards or away from the speaker has a much less damaging effect on the sound though it will change the apparent width of the image and the off axis response of a loudspeaker is notably poorer.

Figure 3.2: A Stereo Sound Reproduction



There have been many attempts to address the limitations of the stereo reproduction system while retaining its elegant simplicity. One of the more practical involved the design of special loudspeakers with a specific polar response that allowed the intensities from each speaker to be correct over a much wider area. [42] The polar output of the pair of speakers was tailored so that as the listener moved to one side the change of angle caused the output of the nearer loudspeaker to effectively fall, and that of the further to rise. This was achieved by mounting the speakers facing vertically downwards onto a specially designed acoustic mirror. Unfortunately this does not deal with the problem of the false phase clues.

It is important to note that the brain is capable of considerable processing, such as the cocktail party effect, where only one voice can be selected from high levels of background noise. With this asset the stereo system can provide good results even with its limitations. Comments like [54] *pp 30* "This explains why a high quality mono recording, made with an accurate microphone in a live acoustic, will contain some clues to directionality even though a singe channel of information is available for reproduction" serve to indicate the power of the brain at reconstructing a scene even without direct information. (A single signal recorded by a normal microphone cannot contain directional information.)

3.2.1 Cinema Sound

In a cinema it is important that the sound image should extend over the entire seating area. The solution used is to record a large number of different tracks and reproduce these from speakers surrounding the listeners. The technology is increasingly becoming available to domestic video users where normally five channels are stored: left, right, centre, surrounds and in addition a LF specific channel is available. The large number of different sources produces a sound image reproduction that, while still far from perfect, produces much better results over a large area.

It is interesting to note that the increasing number of cinema type sound systems in the home and the advent of DVD storage disks may herald the end of stereo's supremacy as a sound reproduction standard. (Digital Versatile Disk or DVD is a storage medium that promises to be able to hold a feature film including five channels of digital sound with ease in a medium the size of a CD. DVD is targeted mainly at the domestic video market but could also be used for bulk data storage.) It seems likely that if such systems become normal that music will be recorded to suit this much more capable standard in the future, possibly with appropriate visual stimulation. At present video tape recordings fail to compete on the grounds of quality, convenience of access, durability and size; all of which are addressed in the DVD standard.

3.3 Human Hearing Abilities and Stereo Reproduction

It is important to understand how the human hearing system functions so that the consequences of practical compromises can be fully understood and imperceptible parts of the image discarded where this eases implementation. Only the aspects of the human hearing system relating to source location are considered here. The ability of the human hearing system to detect small changes in frequency and thus Doppler type distortion is considered separately in section 6.

3.3.1 Location of a Sound in Space

A sound location can be given by azimuth (lateral), elevation (vertical) and distance co-ordinates (essentially the polar coordinates of the sound source regarding the listener as the origin). This representation is convenient as the human sound location ability can be described most consistently in terms of these variables. Considering each in turn starting with azimuth location.

- Azimuth Location

  Of the human sound location abilities the best is for azimuth. Localization becomes more accurate below 1000Hz and above 4000Hz with the smallest errors of the order of 4.6degrees [22]*pp 430*. The Minimum Audible Angle (MAA - the smallest change in source angle that is detectable) is at best 1-2 degrees [22]*pp 432*. The location ability is at its most accurate for low frequency signals (approximately 200Hz) with the source located directly ahead. Poorest localization occurs with the source in front of one ear (90^0 azimuth) and between approximately 1000 and 4000Hz. The location error does not exceed about 20 degrees at any frequency

  There are two mechanisms used to achieve azimuth location, the Interaural Time Difference (ITD) and the Interaural Intensity Difference (IID). The ITD is caused by differing path lengths from the source to the ears and the IID by the acoustic shadow that is cast by the head which reduces the intensity of the sound reaching the shadowed ear. Of these the ITD provides the most accurate location by a small margin. [22]*pp 431* provides a useful plot of location accuracy.

  Human sound localization uses the ITD for location for frequencies up to about 1500Hz [22]*pp 435*, sensitive to about 3^o phase difference. (This mechanism is substantially less accurate by 2kHz and has no sensitivity by 4kHz.) In an alternative test it was demonstrated that ITDs down to about $6\mu s$ can be perceived provided a long pulse is used that contains LF components [22]*pp 435*. The sensitivity to IID is of the order of 1dB across the band though it has been proposed that a thresholding method is used at LF rather than direct intensity comparison [22]*pp 436-437*. At an angle of 30°, the norm for a stereo system, IIDs should be expected to make a significant contribution to the location ability for frequencies of approximately 1800Hz and higher. (Lower frequencies will cause the IIDs to be very small due the size of the head relative to the wavelength.)

For each value of ITD or IID there are two possible solutions for the location of the sound; one is in front a line passing through both ears and the other behind. The accuracy with which we are able to detect the difference is limited, to the extent that when measuring the localization error the location of the sound source was measured either relative to 0^0 or 180^0 azimuth, whichever proved the more accurate [22]*pp 430.* Given that the pinnae (the external parts of the ears) must be responsible for this ability (since the recorded IID and ITDs are identical) and they only have a significant impact at greater then about 4000Hz then the majority of the frequencies used in the experiment in question are too low for this mechanism to be effective and thus this result should be expected. No results have been found indicating the location accuracy of the human hearing using pinael effects only.

- Elevation Location

  A change in the elevation of a sound source does not change the values of either the ITD or IID and thus any ability to locate the elevation of a sound source must be due to pinael effects. As before, aside from noting that these effects exist no record has been found of the accuracy that could be expected has been found.

- Distance Location

  There has been no reference made in the texts consulted to distance perception so it is likely that our ability is entirely due to identifying the sound and then comparing the intensity of the sound with that expected. (The spectral shift due to long distances travelled through the air may also be significant.) This is not true for eyesight where there is a distinct depth perception, although other clues are used as well [16]. Since each eye is able to measure the lateral and elevation angles to an object accurately it is possible for the brain to compare the results obtained and thus obtain some measure of distance. Location by each ear individually is limited to pinael effects and thus it is unlikely to be sufficiently accurate to provide any useful information.

Thus in designing a reproduction system it is most important to satisfy the lateralization ability, followed by the vertical and then the distance. As noted there is no evidence to support that the human hearing is capable of distance measurement and so there is unlikely to be any benefit in this last level of realism. Most audio scenes for reproduction are likely to be arranged on a plane and hence there is unlikely to be any significant amount of vertical information in the direct sound. (There will be reflected sound energy from all directions but the brain will tend not to use this for localization because of time effects.) This leaves the lateral location as being significant.

There are a number systems that are commonly used to reproduce sound, the traditional "stereo" system and at least two different systems designed for use with cinema sound [52].

3.3.2 Time effects and Masking

The abilities of the human hearing system in the time domain are also significant to stereo reproduction. Perhaps the most important effect is the Haas effect, where the first signal to arrive will tend to mask all similar signals for a short period. In [22] masking was found to be a strong effect up to about 15 ms, but was still present at 50 ms. This concurs with the figures given in [50] where it was stated that two similar sounds separated by over about 60-80 ms would be heard as discrete sounds rather than fused into one as would occur below this figure.

It is also interesting to note that the human hearing system can detect a gap in a signal as short as 2-3 ms.

3.3.3 Human Hearing and the Operation of Stereo Reproduction

For the moment the stereo system remains the only mainstream sound reproduction system and thus the more complicated cinema type systems will not be considered in detail.

The simplicity of the stereo reproduction system is its great asset, but it suffers from many limitations. Perhaps the most significant is the small area over which the illusion is effective, the chief justification for the additional complexity of the cinema system. (A partial solution has been offered in the form of the wide imaging stereo technology mentioned previously [42]. This solution is incomplete since it does not address the problem of the the differences in the time of arrival as the listener moves closer to one or other loudspeaker.) The stereo system also suffers from multi-path interference and failure to completely satisfy the human acoustic location ability.

Examining each of the problems in more detail:

- Failure to Satisfy the Human Location Ability

  In the stereo system no attempt is made to preserve the LF phase location clues, those which allow humans to most accurately (by a small margin) locate the source of a sound. Sinclair *et al* [54] state that the phase clues are converted into intensity differences. This is not true with the crossed cardioid microphone recording method since the microphones diaphragms are as close together as is practical and thus the phase difference is parasitic and small. Even if it were true it would be to no avail since with the 30^o angle of a stereo system the human hearing system is incapable of discerning the intensity differences that occur at less than about 1800Hz, as discussed in the previous section.

  Of course in the limiting case, where a signal is produced from one loudspeaker only, the location ability will be completely satisfied. This case will not occur with a true stereo recording, ie. one made with crossed cardioid microphones etc.

  If we consider a signal set somewhat to the left in the stereo image then the left hand ear would perceive high frequencies as being louder. Thus if the signal contained only HF components the location would be placed appropriately. If lower frequency components are used then there will be no intensity difference perceived even if there is an intensity difference at source. The direct signal will arrive at each ear twice, due to the multi-path nature of the system. In each case the first instance will be used for the time based location, even if the second were louder [22]*pp 439*, thus the location perceived will be directly ahead for the LF components. The human location system can also make use of the envelope of LF modulated HF signals for location purposes [22]*pp 438*, but this too will be defeated by the Haas effect. (The Haas effect is that the first copy of a signal to arrive will be used by the brain to establish the location, slightly delayed copies will not be heard individually. This effect is discussed in [22].

  It is also significant to note that even with headphone listening, when IIDs can be produced at low frequencies, it is not possible to balance the effect of an IID with an ITD. The results are, at best, imprecise and the attempt can lead to splitting of the image. Thus supplying phase clues that indicate a central location and IIDs that do not will cause the image to be weak at best and contain spurious doubles at worst. The author of [61] dealt with this problem by only supplying the HF component direct and delaying the remainder of the signal. He theorized that the location would be based on the HF component only because of the Haas effect with the remainder adding colour. The justification of this solution seems a little weak since the first time an LF component arrives it will be identified as unique and will be acted on by the LF location mechanisms in the human auditory system. Sources that produce both HF and LF signals are unlikely to be uniquely or accurately located.

- Multi-path Interference

  The signal reproduced from both of the loudspeakers will reach both of the ears, but taking the case of the left ear, the signal from the right loudspeaker will have to travel further than that from the left; thus there will be the constructive/destructive interference between these largely similar signals. (The signals produced by the loudspeakers would be identical if the sound were to appear in the centre of the stereo image.) The effects can be simulated. It is significant to note that the interaural intensity difference (IID) *never* exceeds 10dB for the 30^o azimuth angle of the stereo system and thus the two signals would be very similar in amplitude. The difference is very small up to about 2kHz and thus this effect is likely to be particularly strong - at 2kHz the wavelength is about 17cm, thus the difference in path length will already be greater than half a wavelength when the most serious cancellation would occur.

3.4 Summary

The stereo reproduction system is the de facto standard for sound reproduction, though this may be set to change with the advent of DVD. The system is simple and elegant but suffers from a relatively poor spatial sound illusion since it is only capable of deceiving one of the human sound location mechanisms. In order for a stereo system to perform well accurate reproduction is required over the large dynamic and frequency range of the human ear.

All of the location information provided by a stereo system is intended to produce IIDs, since ITDs are difficult to reproduce except for with rigidly controlled listening conditions that are impractical for domestic use. The stereo system will not reproduce ITDs when recordings are produced using the method specified by Blumlein [54]. Detectable IIDs will only occur for signals of greater than about 1800Hz for the stereo angle of 30^o thus below this frequency there will be no stereo information with loudspeaker reproduction.

The IIDs are only correctly reproduced over a small area but techniques have been pioneered to reduce this problem by carefully controlling the polar output of the loudspeakers to compensate for changing listening position.

The stereo system suffers from multi-path interference for all signals except those emitted from only one loudspeaker.

Many of the problems with the stereo reproduction system, for instance multi-path interference between the signals from each speaker and limited spatial reproduction, are inherent in the concept and cannot be overcome even with the application of modern techniques. There are, however, a number of areas that could possibly benefit from the application of DSP technology. Designing loudspeakers to fully satisfy the human hearing abilities represents a considerable engineering challenge that would be made easier if closed loop techniques could be used to control the movement of the cone of the loudspeaker. Perhaps the most elegant solution is to make use of the back EMF generated in the voice coil as it moves but some processing will be required if this signal is to be a useful measure. All loudspeakers, but particularly modern ones that are designed to accommodate larger cone excursion and therefore peak velocity, suffer from Doppler distortion which can be reduced using DSP techniques. Signals reflected from the walls of the room where the reproduction takes place are undesirable since they do not exist in the original sound image and are beyond the control of the recording engineer. Any system that is able to reduce or remove the effects of these reflected signals would benefit the reproduction.

If DVD becomes common it is possible that the stereo system will no longer be the default method of reproducing sound. Cinema type sound reproduction systems, that can easily be supported with the many channels DVD offers, offer improvements over the traditional stereo system in many areas. The applications of DSP to improve reproduction already highlighted, however, remain as valid for cinema type systems as they were for the conventional stereo system.

Chapter 4

Details of the Test Equipment and Location

To carry out measurements of loudspeaker operational parameters and room transfer functions equipment was required that was capable of reproducing and recording at least two channels of digital sound at HiFi quality simultaneously.

A block diagram of the test equipment is given in figure 4.1.

The test equipment was based around a SUN sparc 2 which allowed much of the data processing to be carried out using Matlab. Berkeley Camera Engineering, California, provided a card, the SBusDSP-56X, that plugged in to the sparc's expansion bus, the SBus. The card they provided included a Motorola DSP56001 capable of real-time processing independent of the host sparc. (The DSP56001 is capable of only a modest 16 Peak Million Instructions Per Second (Peak MIPS).)

Connected to the SBusDSP-56X was a Legato AES box, a device supplied by Berkeley Camera Engineering that translated the output of the SBusDSP-56X to either the AES of SP/DIF digital audio standard. (The AES standard is intended for professional grade audio equipment and the SP/DIF commercial grade.)

Connected to the AES ports of the Legato AES box were a Legato 18 bit Digital to Analogue Converter (DAC) and a Crystal 18 bit Analogue to Digital Converter (ADC). Some modifications were required to the Berkeley Camera DSP code to allow full use to be made of the 18bit capability of these devices.

A signal source was provided in the form of a CD player that was connected to the Legato AES box's SP/DIF port.

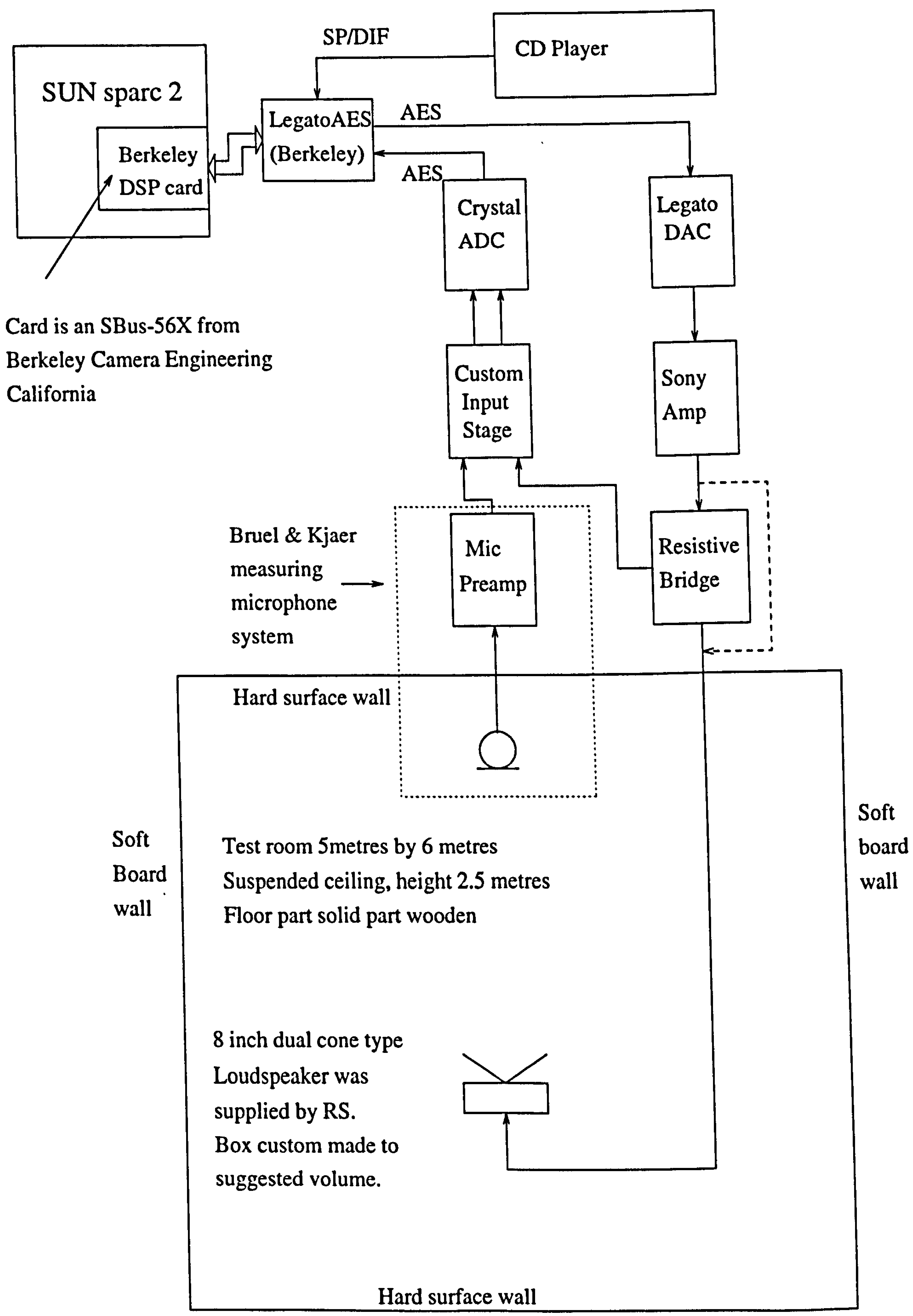
A variable gain stage was custom made for the input to the Crystal ADC evaluation board. The input stage also provided protection and allowed basic signal level monitoring.

The microphone used was a Bruel and Kjaer type 2631 measuring microphone. Due to this microphone's unusual carrier mode of operation it suffered from high internal noise levels. The microphone was, however, accurate even with signals at much less than one a Hertz.

A 100wRMS Sony amplifier took the output of the DAC and drove the test loudspeaker. The loudspeaker used consisted of an 8 inch twin cone driver, sourced from RS, mounted in a basic custom made enclosure. The enclosure was a sealed or infinite baffle type.

Software was written for Matlab to drive the system.

Figure 4.1: Test Equipment Block Diagram



Chapter 5

Feedback Control of Loudspeakers using the back EMF

5.1 Introduction

A moving coil loudspeaker is the most common mechanism used to reproduce sound. The driver consists of a motor coil in a magnetic field attached to a cone which is suspended in a chassis. The driver is mounted in an enclosure. A current passing through the motor coil reacts against the standing magnetic field and exerts a force on the cone. The force causes the cone to move with the displacement controlled by the suspension and the cone's mass. The mechanics are considered in greater detail in the section 5.3.

Careful design has allowed a largely constant Sound Pressure Level (SPL) to be produced for a constant voltage input amplitude over a wide range of frequencies. There are two different types of deviation from the ideal, those due to the concept's inherent limitations and the results of differences between the concept and a practical implementation. It is not difficult to compensate for the limitations of the concept since these are predictable, though impractical excursion may be required of the cone at the lower frequencies. The second class are far more difficult, especially since some of the imperfections may vary between examples of nominally identical drivers or with operating conditions. Despite evolving over many years there is still a great deal of potential for improving the performance of speakers; with the advent of digital storage for musical signals they are probably the weakest link in the HiFi reproduction chain.

HiFi loudspeakers are normally an accepted compromise of conceptual design that best meets the desired specification with all other deviations from the ideal minimized. The professional audio equipment market is different, there are a number of manufacturers that market equalizers for use with their speakers in order to further improve their performance. Even adding fixed equalization cannot correct for all defects in the performance.

Control system techniques can be used to reduce the effects of unpredictable variations. Closed loop control techniques compare the actual output with the desired value and apply a correction on the basis of the difference. A sensor is required to determine the value of the output at any time.

The motion of a coil in a magnetic field causes an EMF to be developed across it that is dependent on the velocity. The voltage is called the back EMF and is generated across the voice coil of a loudspeaker when the cone is in motion and can be used to sense the coil's velocity. If the coil has a current passing through it in order to make it move then the voltage created by it's motion is still present. It it is more difficult to measure this EMF in a real system where the coil will have a real resistance since the driving current will also cause a potential to appear across the coil. If a resistive bridge is used the effects of the DC resistance of the coil can be compensated for and the bridge output will yield the back EMF only. If the coil's resistance changes then unless the bridge is re-balanced the bridge will yield spurious results.

The concept researched here is to make use of this back EMF to sense the cone's velocity in order to allow the use of feedback control to eliminate the effects of the more unpredictable imperfections of a real driver. (Note that feedback control will inherently tend to linearize all imperfections in the response of the system, provided that the feedback signal is linear.)

5.2 Literature Survey

A commercial system making use of motional feedback was marketed by Philips but was discontinued in the early 1980s. When contacted Philips would not provide details of this device. No external references to this system have been found. It is highly likely that this system made use of a piezo accelerometer to provide cone motion information.

The concept of using the back EMF to allow feedback control of loudspeaker drivers has been investigated previously as a method of of controlling the cone damping [67]. The application of the technique to improving the frequency response of the device has been considered [69] [68]. The technique has also been studied at Bangor University in the past [38], [29] and [4] but due to problems the method was discarded by [31] in favour of a piezo acceleration sensor. The principal problem reported was the effects of the series resonance between the voice coil's self inductance and the cone mass. Attempts to reduce this effect with a filter after the bridge apparently were unsuccessful due to the modulation of the voice coil's self inductance which was reported to change by up to 10% at large excursions.

Other methods of feedback control were also developed, some in detail since they demonstrated promise. Examples typically either made use of a piezo electric sensor to measure the cone acceleration, as investigated in Bangor, or used electrostatic principles to measure the cone position [12]. The latter method had the advantage that it measured the average cone position rather than the motion of the throat, which can be quite different above the first cone break up mode.

5.3 Subject Revision

5.3.1 The Moving Coil Loudspeaker

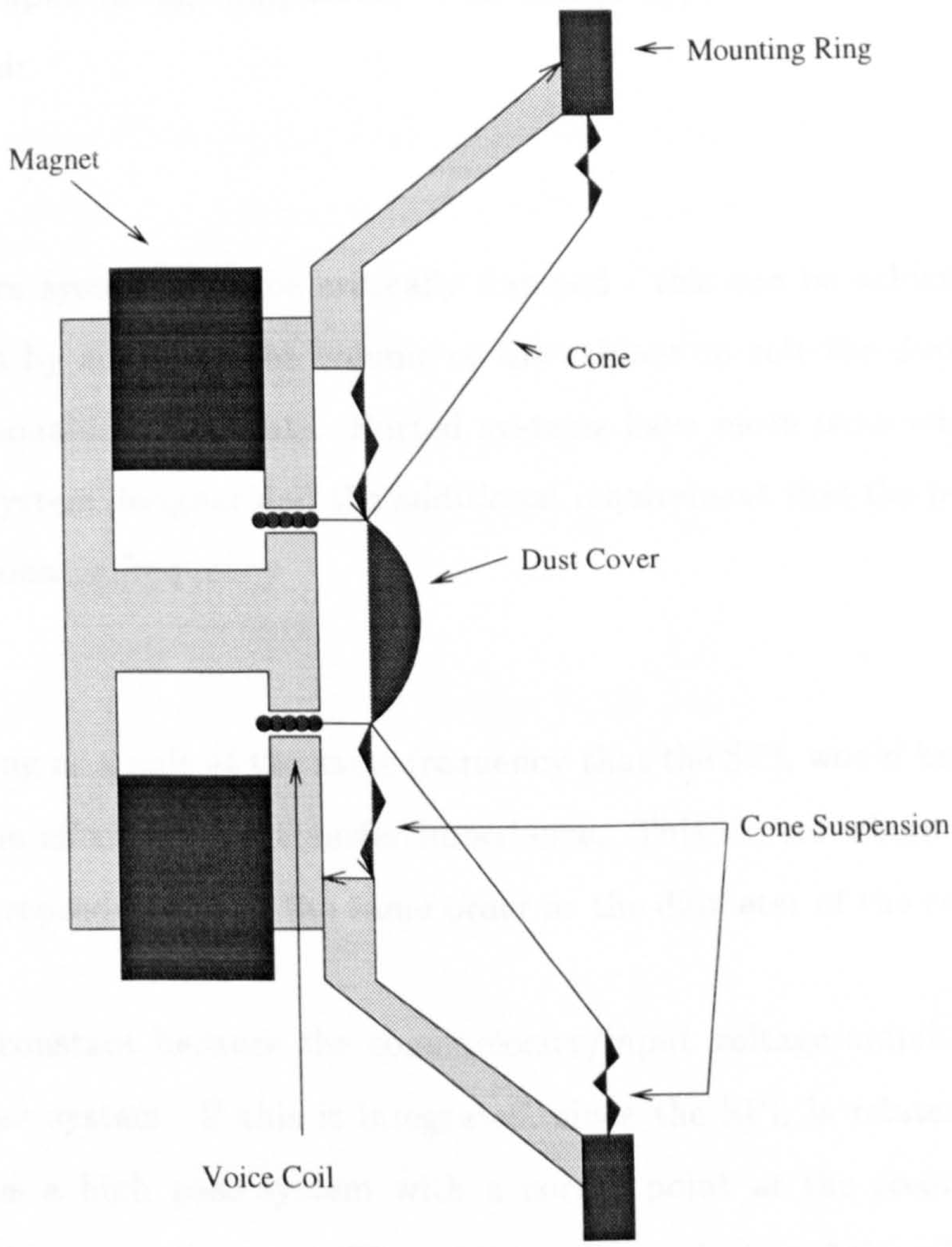
Concept

A moving coil loudspeaker is a device that is designed to translate an electrical analogue of a sound into the appropriate pressure waves in the air. A loudspeaker consists of a driver in an enclosure. The enclosure is an important part of the system since it modifies the output polar diagram and effective bulk parameters of a driver.

The moving coil driver consists of a cone suspended in a chassis. At the throat of the cone a coil is attached, the voice coil, that is immersed in a magnetic field. A cross-section of a typical driver is given in *figure 5.1.* When an electrical current is passed through the voice coil it exerts a force against the standing field which causes the cone to move. When the cone is moving in the magnetic field there is a voltage across the coil and since the cone is being driven this is reverse polarized, the back EMF. In a perfect system where the voice coil had no resistance, the back EMF would exactly equal the driving EMF.

Practical loudspeakers voice coils have a resistance of the order of 8Ω and are inefficient, the bulk of the wasted energy is dissipated in the voice coil. Since ideally the voice coil needs to be as light as possible it is difficult to envisage a system that could keep its temperature roughly constant. The materials used to make the voice coil, typically copper or aluminium, change resistivity with temperature and thus the resistance of the voice coil must be expected to change during the operation of the device. These ideas are discussed in more detail in [6].

Figure 5.1: A Moving Coil Loudspeaker Driver



The driver is placed in an enclosure mostly to control the radiation from the back of the cone. (If this is not done, as in the case of an open-backed enclosure, the bass response would be severely compromised by the rear pressure wave bleeding round to the front.) There are two different schemes that could be followed, either the enclosure can be completely sealed (which is called an infinite baffle after the similarity of the results obtained to the mathematical simplification) or there can be a carefully designed vent or port. The port is designed to ensure that the radiation from the back of the cone is emitted in phase with that from the front at frequencies a little below the resonant frequency of the driver cone, etc which has the effect of extending the bass response. The system can be thought of as a Helmholtz resonator whose resonant frequency is a little below that of the driver/enclosure system. Ported enclosures offer significantly better bass response for a given driver and enclosure size but at the price of poorer impulse performance and faster initial LF roll off. Ported enclosures are used almost exclusively for modern miniature HiFi loudspeakers.

Modern amplifiers are typically designed to have a very low output impedance, high input impedance and a flat frequency response. Loudspeakers, therefore, are required ideally to produce a constant SPL with input voltage amplitude. This occurs approximately above the resonant frequency provided:

- Critical Damping

  The loudspeaker/enclosure system must be critically damped - this can be achieved in a infinite baffle system by adjusting the volume of the cabinet to suit the design of the driver, within reasonable constraints. Ported systems have more parameters under the control of the system designer and the additional requirement that the port resonance is set to a reasonable frequency.

- Cone Breakup

  The cone must stop moving as a unit at the same frequency that the SPL would begin to increase because of the effects of the transfer impedance. This occurs where the wavelength of the sound reproduced is of the same order as the diameter of the cone.

The output is approximately constant because the cone velocity/input voltage amplitude forms a second order band-pass system. If this is integrated, since the SPL is related to acceleration, then the result is a high pass system with a corner point at the resonant frequency - or an approximately constant output. The additional complexity of the effects of changing transfer impedance are offset typically by the cone ceasing to move as a single entity.

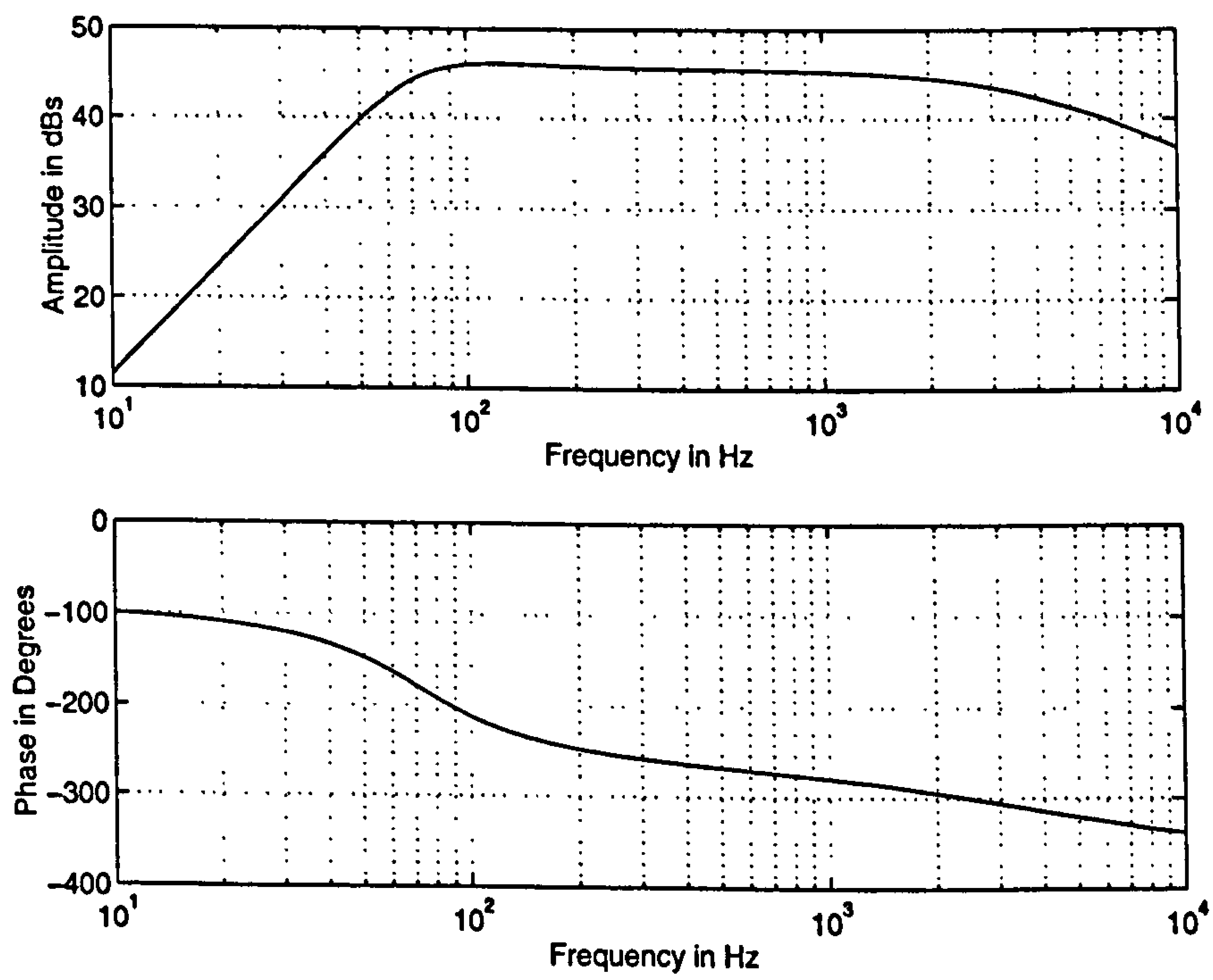
It has been noted that the temperature of the voice coil is expected to change during the operation of the loudspeaker. Since the voice coil of a speaker is typically made of a metal (aluminium or copper) the resistance of the coil will rise with its temperature. When such a speaker is driven from a constant voltage amplitude source the current flowing through the coil will decrease as the temperature of the coil rises and thus the force driving the cone will decrease and with it the acoustic output. It is therefore essential that the temperature of the cone is maintained if the acoustic output is to be constant - which is unlikely to occur due to the transient nature of music signals. This limitation is normally ignored although there are possible solutions, such as making use of constant current amplifiers or amplifiers with variable negative output impedance [6]. (Note that constant current drive does not result directly in constant SPL in the same way that constant voltage drive does and thus some additional filtering/control is thus required.)

Limitations and Methods for Improving Performance

The drivers that are used to make modern loudspeakers are a great improvement on than their predecessors. The improvements are mainly due to fine tuning a concept that has remained largely unchanged for years. Despite this and especially with the advent of digital media the loudspeaker is probably the weakest element in the reproduction chain.

The theoretical output of an infinite baffle system according to the conceptual model is given in *figure 5.2* for the test speaker. It can be seen that the roll off at low frequencies is 40dB/decade.

Figure 5.2: Theoretical SPL Output of an Infinite Baffle Speaker



Ported systems typically have a faster initial roll off since the system is of higher order. An initial roll off value of 60dB /decade or 18 dB/octave is expected since these are third order systems as opposed to the sealed box type which are second order. (During the initial roll-off the output of the port is still dominant and it's nominal velocity rolls off at 40dB/decade, since this is below the passband of effectively two band-pass filters in series. The acceleration rolls off at an additional 20dB/decade and thus the SPL rolls off at 60dB/decade.

More details can be found in [12]*pp 92*, [18].

It is possible to compensate for the limited LF performance by adding a filter to reverse these effects. In practice such correction may not be desirable since the cone displacement required to attain the extended LF performance rapidly becomes excessive. (The reason is that large excursions are required to produce the SPL at LF, any system to boost the LF performance electronically will suffer from this problem.) The excursion problem is a fundamental limit; to maintain an SPL the peak cone acceleration must remain constant which leads to ever increasing excursion as the frequency decreases. The larger the cone the smaller the excursion for a given frequency and SPL, thus large cones are used for high power bass reproduction. Ported systems have an even faster LF roll off and thus are even more difficult to enhance in this way. (In addition the power required to attain these very large displacements rises very rapidly indeed.) The only systems that will undeniably benefit from these response extension techniques are heavily over-damped infinite baffle systems. In this case the SPL falls quite slowly initially and therefore the amount of correction required is reasonable provided that extension in bandwidth attempted is not too great. It should be noted that rapid changes in the loudspeaker response, for instance sharp peaks in the amplitude-frequency response or rapid changes in the phase-frequency response, are considered subjectively inferior to changes of the same magnitude that occur more slowly with frequency.

The LF limitation can be expressed in the following equation [12]

$$W_{max} = k_p(Ax_{pk})^2 f_3^4 \tag{5.1}$$

where

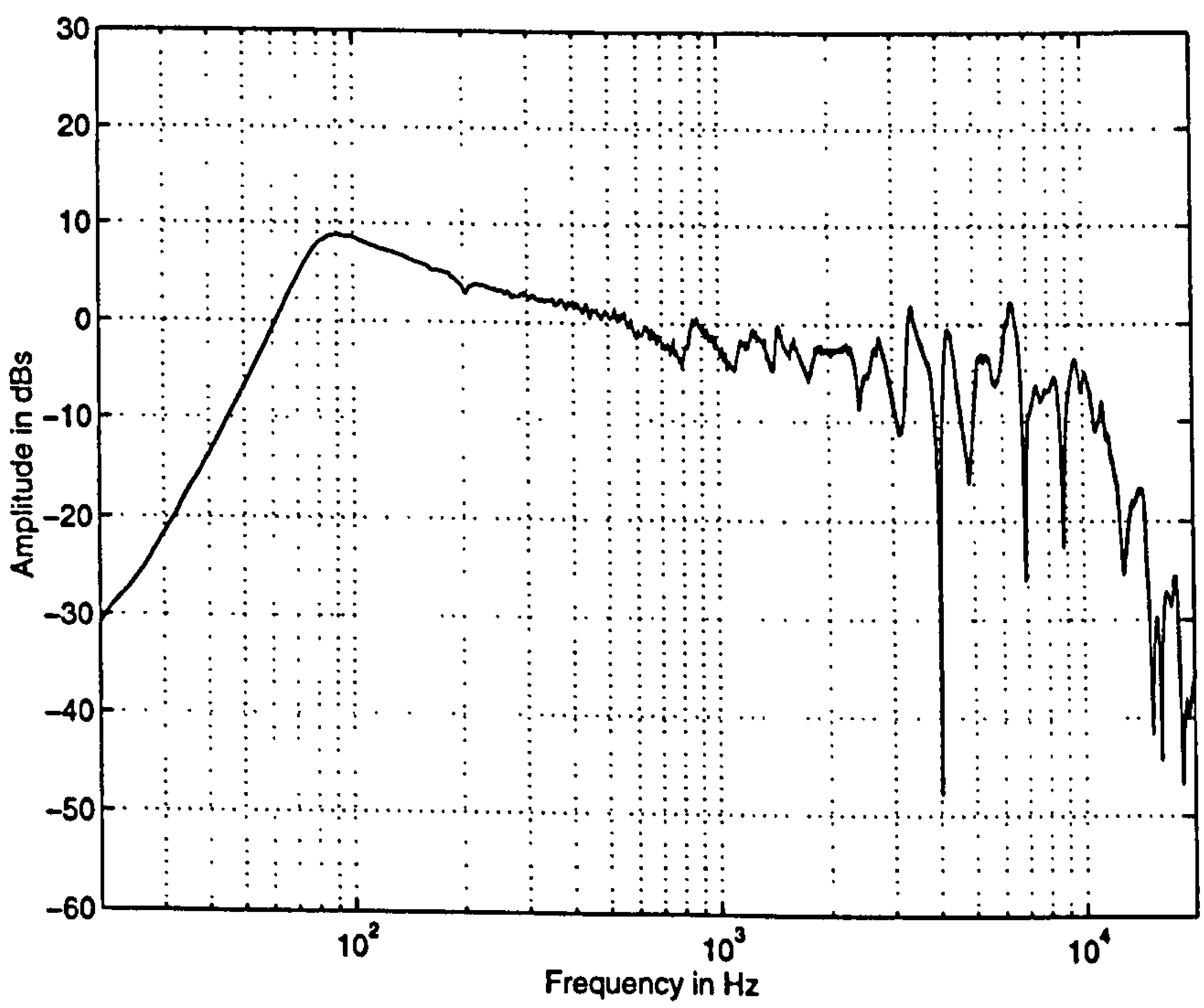
W_{max}	=	Maximum power output
k_p	=	power constant
A	=	Piston area
x_{pk}	=	Peak excursion
f_3	=	-3dB LF roll-off point

This expression states mathematically the fundamental physical limitations of a loudspeaker. Increasing the power output required at the roll off point inherently requires that the excursion increase according to this expression etc.

It is important to note that the frequency response of the loudspeaker is considered in a free-field environment typically, but in practice a typical listening room will have a considerable impact on the performance obtained. In practice a perfectly flat LF response may be undesirable since the reflections from the walls will substantially increase the effective output of the loudspeaker.

To summarize, it is possible to extend the bass response, but it is only likely to be practical in limited circumstances. The real advantage of applying correction lies in the fact that real speakers are far from able to produce the "ideal" response given above, there are many other small defects and colourations in the output. Using the test speaker and the measuring microphone positioned on-axis at a distance of approximately 0.5m (as a compromise between avoiding near field and room acoustic affects) the amplitude response of the test speaker was measured. The response of the speaker is given in *figure 5.3.* Note that the amplitude scale has been arbitrarily normalized. This speaker is not particularly high quality but the defects suffered by a good grade speaker are expected to be similar in appearance if not degree.

Figure 5.3: SPL Output of the Test Speaker in Sealed Enclosure

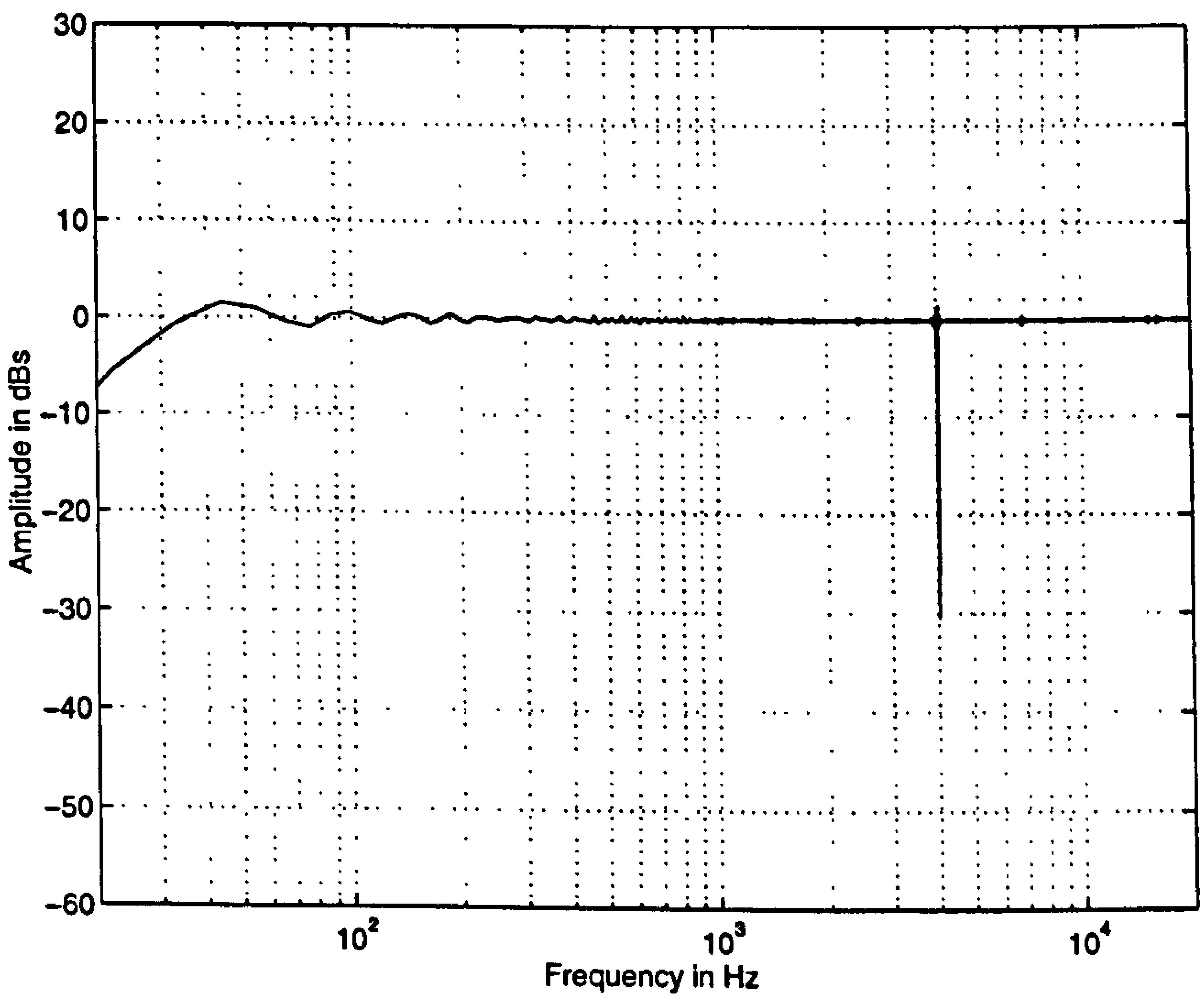


The fluctuation in frequency response at the higher end of the band becomes extreme as the cone enters break-up modes. (Where the cone is not moving as a single unit.) Higher quality systems invariably have multiple drivers to avoid this extreme colouration and gradual reduction of output.

There are essentially two approaches that can be taken when dealing with a system such as a loudspeaker, either a fixed filter can be inserted into the signal path which is carefully designed to make good the defects in performance or closed loop control techniques can be used; of course both methods could be used in combination.

If open loop adaptive techniques similar to those that will be used in section 7 were employed to correct the speaker's response alone the improvement in the output SPL that could be achieved is indicated in figure 5.4. (Note that the improvement is simulated against a model of the loudspeaker that takes no account of any non-linear effects. For this reason and other practical difficulties the improvements would not reach the standard indicated in the graph.) It is possible to arrange for this type of filter to adapt to compensate for *slowly* changing system parameters provided that a suitable performance measure such as a microphone is available.

Figure 5.4: SPL Output of the Test Speaker in Sealed Enclosure



A real loudspeaker will also suffer from non-linear imperfections. It is possible to design an open loop filter that can deal with some non-linear imperfections but the analysis becomes more difficult. There have recently been a number of methods developed that are reported to yield good results, [60] and [20]. Despite these recent improvements closed loop techniques remain more flexible and able to cope with a wider range of response imperfections.

Despite the limitations of open loop correction it remains the only method commonly seen in commercial systems.

Other Types of Driver

There are other types of driver in production but none are as common as the moving coil type. Most alternative types are either based on electrostatic attraction/repulsion or piezo electrical effects. Electrostatics have very limited diaphragm travel and thus need to be large to achieve a reasonable SPL or to be combined with a moving coil device to reproduce the bass frequencies. Electrostatics theoretically have significant advantages as well, the most significant of which is the diaphragm used can be very light as the force is distributed evenly which allows detailed, uncoloured reproduction; there are a number of high end systems on the HiFi market, for instance the Quad ESL.

Piezo effect drivers are common but in contrast to the electrostatics are rarely, if ever, used for anything other than the very cheapest systems. The units are typically cheap to manufacture but the performance that can be achieved seems to fall below that demanded of modern HiFi speakers. Piezo effect devices are only used for the HF drivers in multi-way systems due to their limited travel.

There is no evidence to suggest that either of the above, or any emerging technology, will displace the moving coil loudspeaker as the default standard design for HiFi application, particularly for the low frequency part of the audio spectrum.

5.3.2 Loudspeaker modelling

An important stage in developing any control system is to develop a mathematical model of the system to be controlled. To be of value the model must be simple enough to manipulate and yet maintain sufficient accuracy so as not to be misleading. The model used has been the default standard for many years and thus, while it may not give excellent results, its limitations are well known. As an exercise the accuracy of this model was checked in the lab and the results were in good agreement with the predictions. Theil and Small are usually attributed with the research leading to this model (for instance [55] [58]), to the extent that the detailed parameters by which Loudspeaker manufacturers select drivers have become known as "Theil Small" parameters.

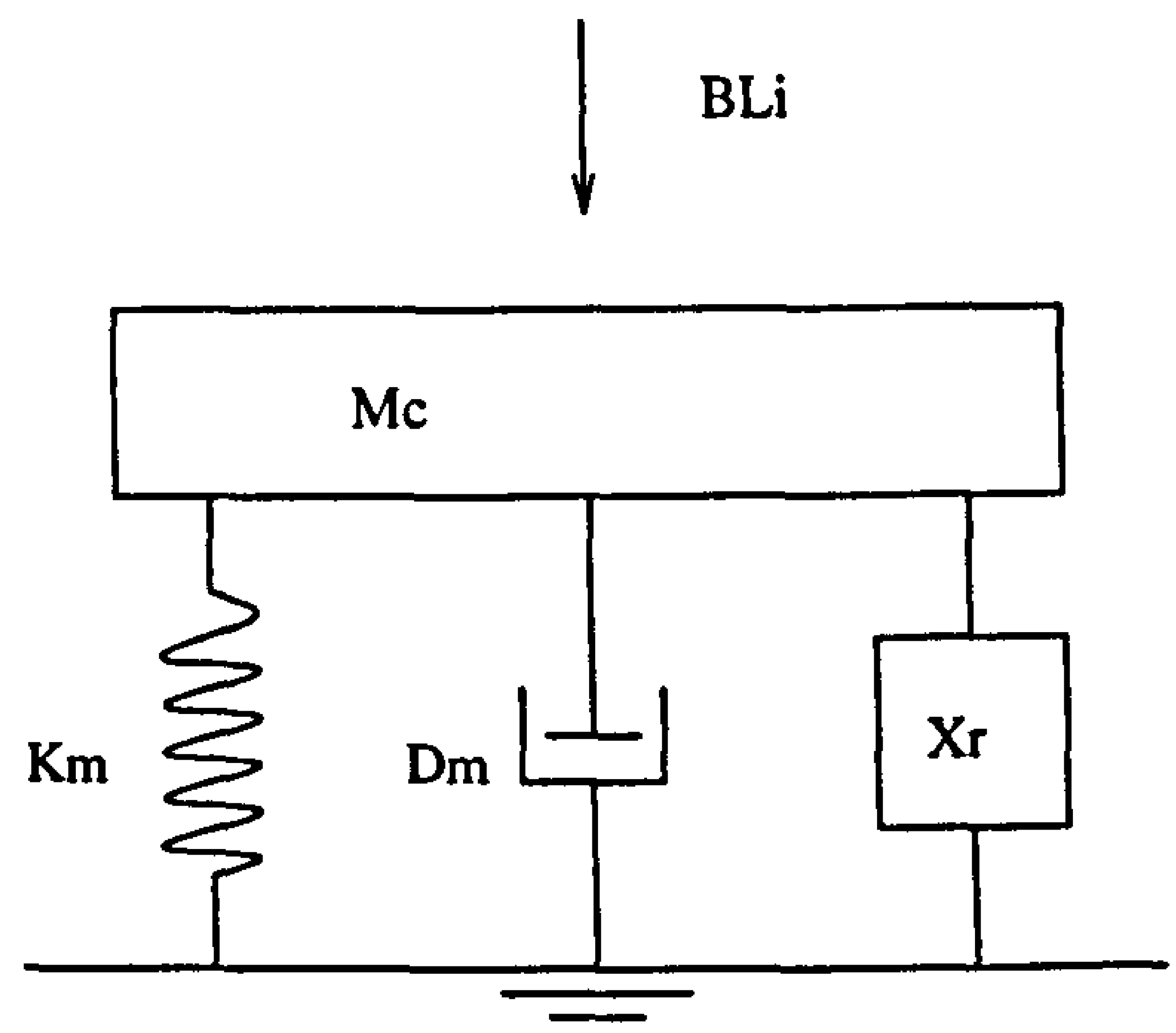
A loudspeaker can be best described as a device that converts a voltage input into a SPL. The driver can be considered as a device that converts a voltage into a cone velocity, though in practice the enclosure will affect the cone velocity. The simplest type of enclosure to model mathematically is the sealed box and thus this will be used for most of the analysis. An sealed box type enclosure may be thought of as an additional restoring force on the cone (due to the air trapped in the enclosure) and an air load mass. (The acoustic impedance may be modelled in more detail as is discussed later.)

At low frequencies it is sufficient to use bulk parameter models but at higher frequencies these simple models become increasingly inaccurate. There are two different effects that contribute strongly to the degradation, the cone ceases to move as a single surface after the first breakup mode and it is no longer sufficient to consider the distributed parameters as a single bulk. Dealing with either of these limitations is problematic.

Since a driver operates both in the electrical and mechanical domain it is common to use one of the many schemes that allow the entire system to be represented in a single domain. For this analysis electrical analogues have been used though any other could have been chosen without change to the final result. The transformation between domains is handled by a transformer. The rational behind the scheme chosen is that all "through" variables can be considered similar, as can all "across" variables; thus voltage is similar to velocity and force to current. Other schemes use quite different rationales to the same end.

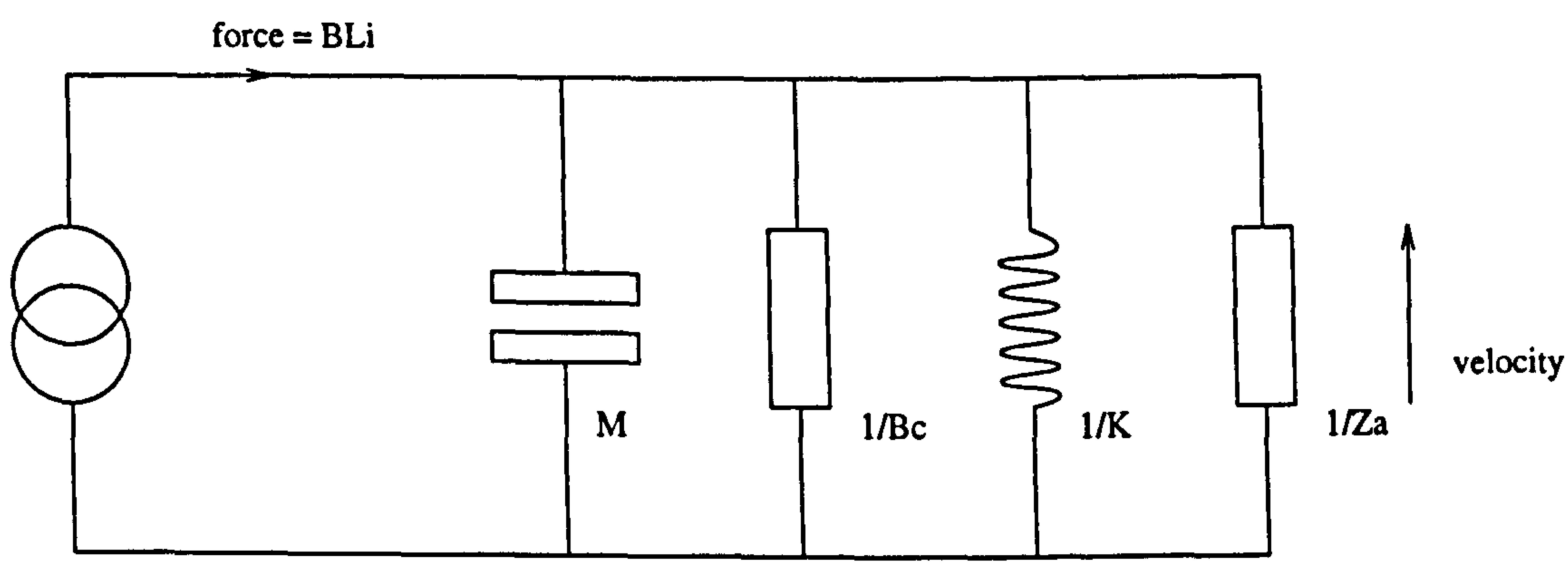
Starting with the mechanical part of the driver; a force BLi drives the cone mass M_c whose motion is controlled by a compliance K_m and a translational damper B_c which result from the cone suspension. In addition to this there is a radiation impedance for the cone/air interface.

Figure 5.5: A Mechanical Model of a Moving Coil Driver



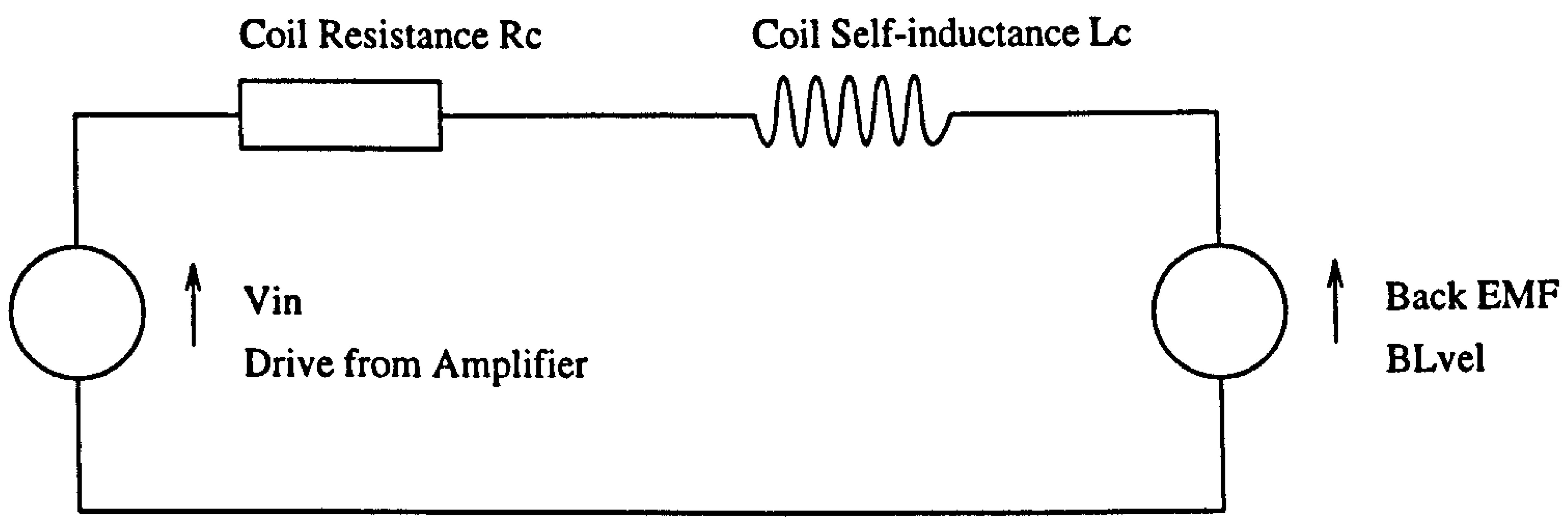
When rendered into electrical analogues according to [15] the resulting circuit is given in *figure 5.6.*

Figure 5.6: An Equivalent Circuit for the Mechanical Aspects of a Driver



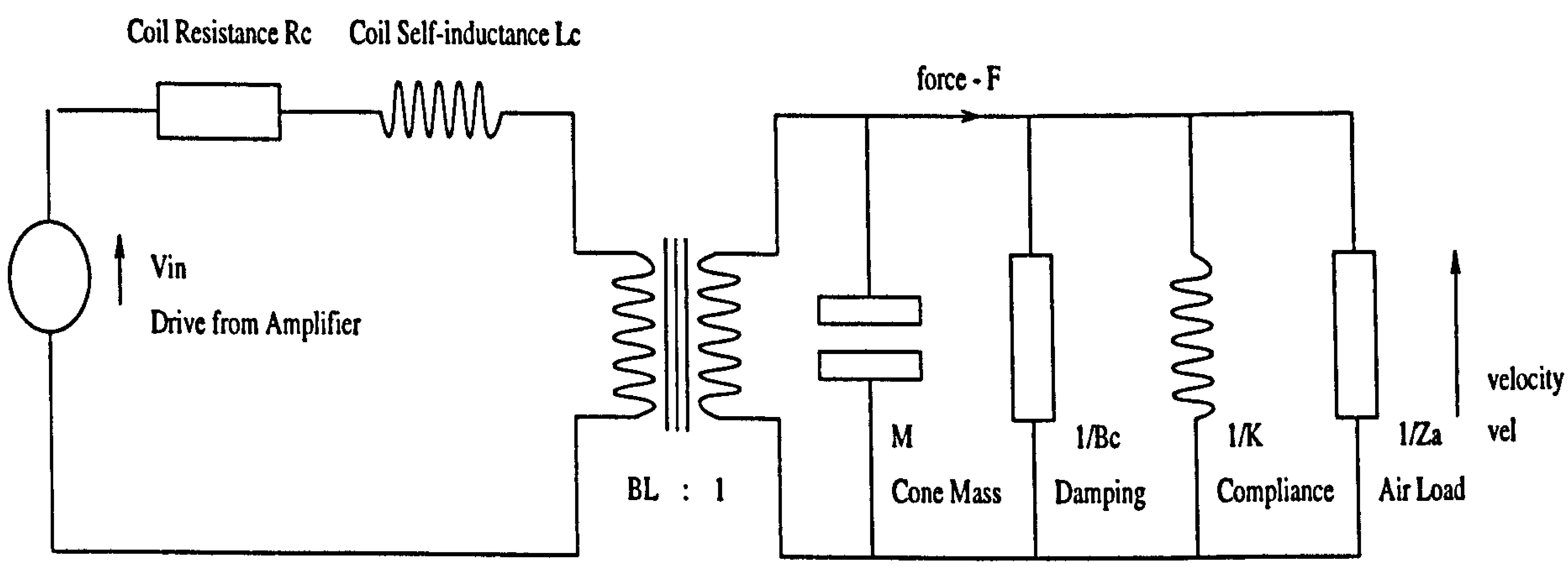
The electrical part of the driver is easily represented as an approximate equivalent circuit. It is important to note that the bulk parameters presented here are, in reality, distributed. At low frequencies this degree of approximation is not significant.

Figure 5.7: An Equivalent Circuit for the Electrical Aspects of a Driver



These two circuits may be coupled by making use of a transformer since F is equal to BLi and vel to V_{in}/BL.

Figure 5.8: The Complete Equivalent Circuit for a Driver



Using this model and standard circuit analysis techniques it is easy to derive the transfer function for $\frac{vel}{V_{in}}$, where vel is the cone velocity and V_{in} is the input voltage.

Firstly the impedance of the series and parallel elements considered independently.:

$$ser = R_c + L_c s \tag{5.2}$$

$$par = \frac{1}{MsB_c + \frac{K}{s} + Z_{a(s)}} \tag{5.3}$$

Using transformer theory the velocity can be expressed as:

$$vel = \frac{\frac{V_{in}}{BL}par}{\frac{ser}{B^2L^2} + par} \tag{5.4}$$

$$= \frac{V_{in}BL}{\frac{ser}{par} + B^2L^2} \tag{5.5}$$

This may be expressed as a transfer function:

$$\frac{vel}{V_{in}} = \frac{BL}{\frac{ser}{par} + B^2L^2} \tag{5.6}$$

Making the obvious substitution:

$$\frac{vel}{V_{in}} = \frac{\frac{BL}{R_cK}s}{\frac{L_cM}{R_cK}s^3 + \left[\frac{L_c(B_c+Z_a)}{R_cK} + \frac{M}{K}\right]s^2 + \left[\frac{B_c+Z_a}{K} + \frac{B^2L^2}{R_cK} + \frac{L_c}{R_c}\right]s + 1} \tag{5.7}$$

Other functions, for instance the impedance, may be derived by a similar method. In practice the analysis can also be performed numerically by programs such as Matlab, which reduces the analytic effort.

5.3.3 Acoustics Relating to Loudspeakers

A detailed treatment of the acoustics relating to loudspeakers is beyond the scope of this document. There are two principal results that are required for the modelling, how the acoustic pressure relates to the motion of the loudspeaker's cone and the resultant force due to the motion that the cone experiences. The first of these is the most important since it is essential that the acoustic pressure is simply related to the cone movement or knowledge of its motion will be of no value.

The acoustic pressure, and indeed the acoustic impedance, are normally calculated for flat pistons. (A cone would be far more difficult to deal with mathematically than a piston.) This is a tolerable approximation while the wavelengths remain comparable to the cone radius but at high frequencies it becomes questionable. (Note that for a 200mm driver the cone radius becomes comparable to the wave length at 3.4kHz, at this type of frequency much of the modelling is of questionable accuracy.) The majority of acoustics texts, such as [32]*pp 169*, produce the expression for the acoustic pressure to be expected due to the motion of a piston in an infinite baffle:

$$\mathrm{p} = \frac{j\rho_0 ck}{2\pi r}Q_p e^{j(\omega t - kr)}\left[\frac{2J_1(ka\sin\theta)}{ka\sin\theta}\right] \tag{5.8}$$

where J_1 is a Bessel function which may be expanded as a power series according to:

$$J_1(x) = \frac{x}{2} - \frac{x^3}{2^2.4} + \frac{x^5}{2^2.4^2.6} - \frac{x^7}{2^2.4^2.6^2.8} \cdots \tag{5.9}$$

and

p	=	acoustic pressure
ρ	=	density of transmission media
c	=	speed of sound in air
k	=	wave number
r	=	radial distance of point from centre of piston
θ	=	angle between a line from the centre of the piston to the point and its axis
Q_p	=	source strength

On axis θ is zero and therefore the directivity function, that in square brackets in 5.8, is unity. In this case the expression reduces to that of a hemisphere in an infinite baffle.

This analysis is subject to the limitation that the distance of the point (r, θ) to the centre of the piston is large compared to the radius of the piston.

The problem is that this type of analysis uses harmonic solutions and thus the answer cannot be used directly for a general signal. (It is possible to plot the results for many sine waves of different frequencies and thus establish the frequency response.)

An alternative analysis can be found in [40]*pp 387*. Here the analysis is for a general signal.

$$p = \frac{\rho c^2}{4r} A_{(t-\frac{r}{c})} \qquad \theta = 0 \tag{5.10}$$

where A represents the piston acceleration.

This result is limited to far fields only, as with the previous method. From this result it may be deduced that the acoustic pressure on axis is directly related to the piston acceleration.

A general expression for the acoustic impedance is desirable since the energy dissipated in air will affect the motion of a loudspeaker's cone. Unfortunately the only analyses that have be found rely on harmonic solutions and thus cannot be used directly in a system transfer function. (Note that it is possible to make a first order approximation since the radiation reactance is always positive. The effect of this impedance is therefore equivalent to adding to the mass of the piston an additional mass given by: $m_r = a^3 \rho_0 \frac{8k}{3}$ for $2ka < 1$ or approximately 300Hz for a 200mm driver[32]*pp 181*.)

If harmonic solutions are used in the analysis the result obtained is [32]:

$$Z_r = \rho_0 c \pi a^2 \left[R_1(2ka) + jX_1(2ka) \right] \tag{5.11}$$

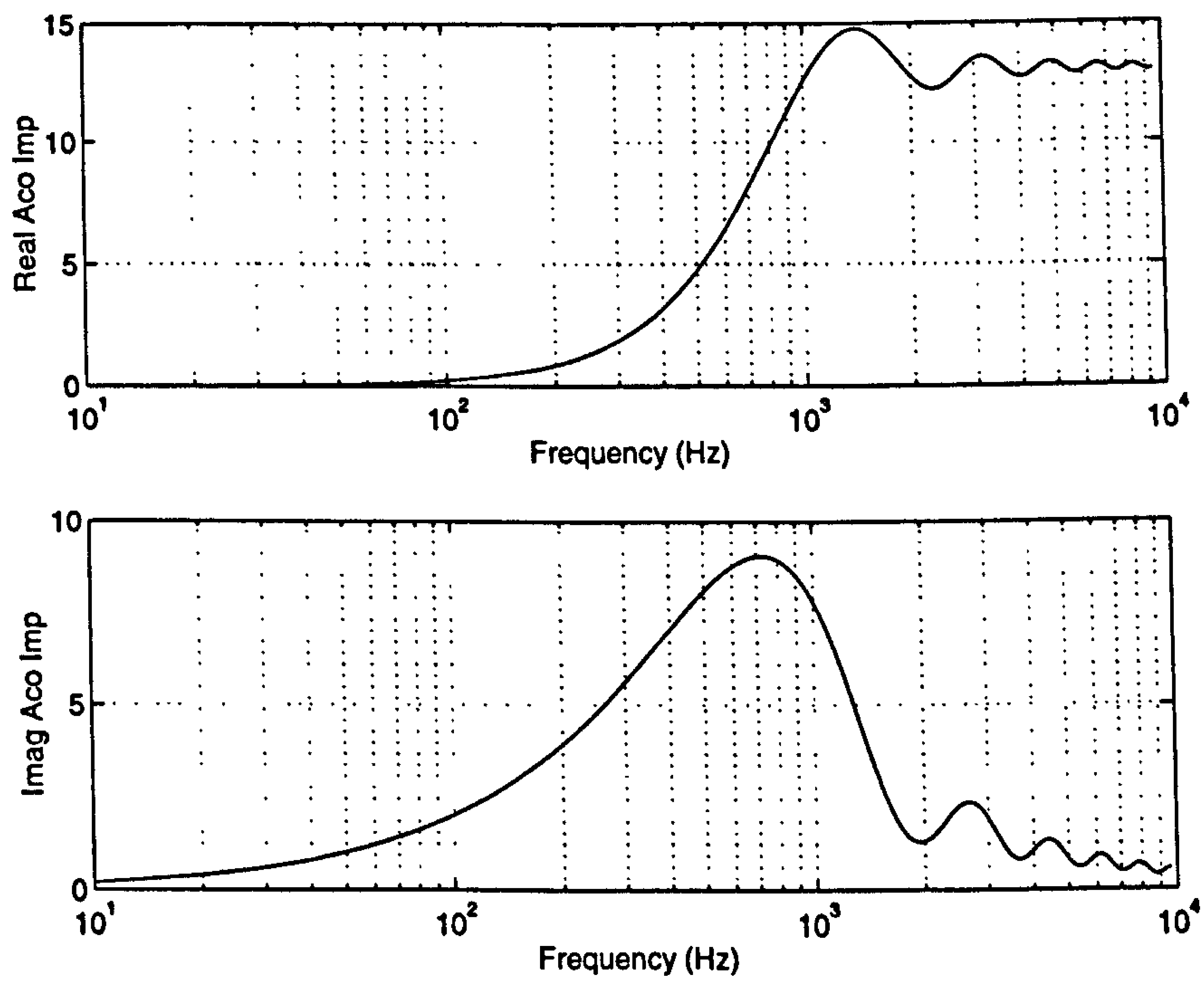
where:

$$R1_{(x)} = \frac{x^2}{2.4} - \frac{x^4}{2.4^2.6} + \frac{x^6}{2.4^2.6^2.8} - \cdots \tag{5.12}$$

$$X1_{(x)} = \frac{4}{\pi}\left(\frac{x}{3} - \frac{x^3}{3^2.5} + \frac{x^5}{3^2.5^2.7} - \cdots\right) \tag{5.13}$$

This result is expressed graphically for the test speaker in *figure 5.9*

Figure 5.9: Acoustic Impedance of a Piston in an Infinite Baffle



5.3.4 Cone Movement Sensors

There are a number of different methods that can be used to measure the movement of the cone of a driver.

- Piezo Sensor

  This is probably the most common sensing method used for applying negative feedback to Loudspeakers. It is relatively cheap to apply, has reasonable noise immunity and need not be too detrimental to the performance of the speaker.

  This method relies on the sensitivity of piezo-electric materials to physical strain. A disk of the material is placed between the cone and a small mass, when the cone accelerates it exerts a force through the piezo-electric material to the mass since these are all rigidly connected. In response to this the piezo electric material produces a charge. This output can be used to determine the cone's acceleration and therefore can be used as a feedback signal to control it's motion.

According to [31] the problems encountered were principally resonances in the sensor construction and the effects of the additional mass. In order to reduce the effects of resonance to reasonable levels the sensor had to be made unreasonably massive. Despite these problems it is likely that the concept would work well provided that the sensor assembly was included at the time of the driver's manufacture.

Since this method only senses the cone's movement where it is attached it can give no useful indication of motion beyond the first break-up mode.

- Capacitive method

  The cone of the driver(s) is typically metallized and a conducting grid is placed in front of it forming a capacitance that will change with the position of the cone. This arrangement may be used either as a condenser microphone, where the signal is obtained directly [12]*pp103*, or by placing the capacitance in an AC bridge. The latter may give better results at LF.

  This method is unique in that it will give an averaged result over the entire cone and thus will take into account the effects of cone break-up. (Of course for a system to function this way the grid would have to be a constant distance from the cone over its area.)

- Back EMF from voice coil

  Any coil moving in a magnetic field will have a potential across it related to the physical layout of the system, the field strength and the velocity of the coil. This EMF is always present, even in driven systems. In a real driven system it is not possible to simply measure the back EMF as the voltage appearing across the coil as there will be a voltage component due to the drive current. If the coil is included in a resistive bridge these effects can be removed.

  If this back EMF can be measured accurately then the velocity of the coil can be calculated which is sufficient for feedback control of a driver. In practice additional problems are encountered due to the self inductance of the coil and parasitic effects.

5.3.5 Current Methods of Compensating for Speaker System Imperfections

There are a number of manufacturers that have systems on the market making use of fixed digital filters to compensate for system imperfections, typically in the professional audio market. The simplest types of filters will not be able to compensate for any non-linear imperfections but there has been a number of papers claiming more advanced filter concepts especially suited to dealing with non-linear imperfections in speaker system operation [20] and [60].

At present there are no bulk manufacturers producing units with any form of feedback control. As noted previously Philips produced a unit in the early 1980s but were unable to provide further information concerning the product.

5.3.6 Closed Loop Control Techniques

The performance of many types of system can be be improved by the application of negative feedback; this is the most common form of closed loop control and a concept that has been used in control and electronics application for many years. The output achieved is compared with the desired value on a continuous basis and a correction signal is fed to the input. For many types of system application of these ideas lead to a dramatic improvement. This technique can be equally efficacious against linear and non-linear defects in performance.

Closed loop techniques rely on having some measure of the final performance of the system; perhaps the most obvious choice for a loudspeaker system would be a microphone arranged in front of the system. In practice there are problems with this method, perhaps the most serious is the delay between the movement of the loudspeaker cone and the sound reaching the microphone which will cause stability problems; moving the microphone closer to the cone to minimize this element will move the microphone into the near field which is not necessarily simply related to the the output SPL. In addition outside the near field the effects of the room acoustic are probably significant not to mention other methods are more desirable from a purely practical stand-point. There have been many methods investigated over the years. One of the most desirable concepts is to make use of the back EMF, which is generated when a coil is moving in a magnetic field, to sense the cone velocity. The chief advantage of this method being that additional parts or unusual modifications are not required to the driver.

A Loudspeaker includes the following that may change over time or with prevalent conditions.

- Magnetic flux density

  It is unreasonable to assume that the magnetic flux density will remain constant over the entire excursion of the coil or as the driver ages.

- Cone Mass

  The cone mass may change by absorption of water vapour etc.

- Suspension Stiffness

  The suspension stiffness has been reported to change significantly with changes in the relative humidity and temperature.

- Damping

  The damping will also depend on the cone suspension but in addition eddy currents in the magnet will also have an effect.

- Coil Resistance

  The resistance of the voice coil should be expected to change by approaching a factor of two over its operating temperature range [6]. (Maximum operating temperatures of over 200°C are possible with some modern loudspeakers.) *Birt* also indicated that the voice coil temperature time-constant should be expected to be in the range of 1s for a tweeter rising to about 40s for a large woofer. *Birt* noted that changes in the coil resistance are reflected in the Q of the amplifier/speaker combination and thus affect not only the sensitivity but also the frequency response. The effects are complicated yet further if a passive crossover is used.

Changing any these parameters will alter the performance of the speaker. Every attempt is made at the design and manufacture stages to produce a device that is stable and consistent in operation but inherently there are compromises. In addition the performance of the loudspeaker will not be perfect even at manufacture. It is possible to insert a filter designed at the manufacturing stage to compensate for those problems evident initially but this cannot take into account changing operating conditions or aging of the device.

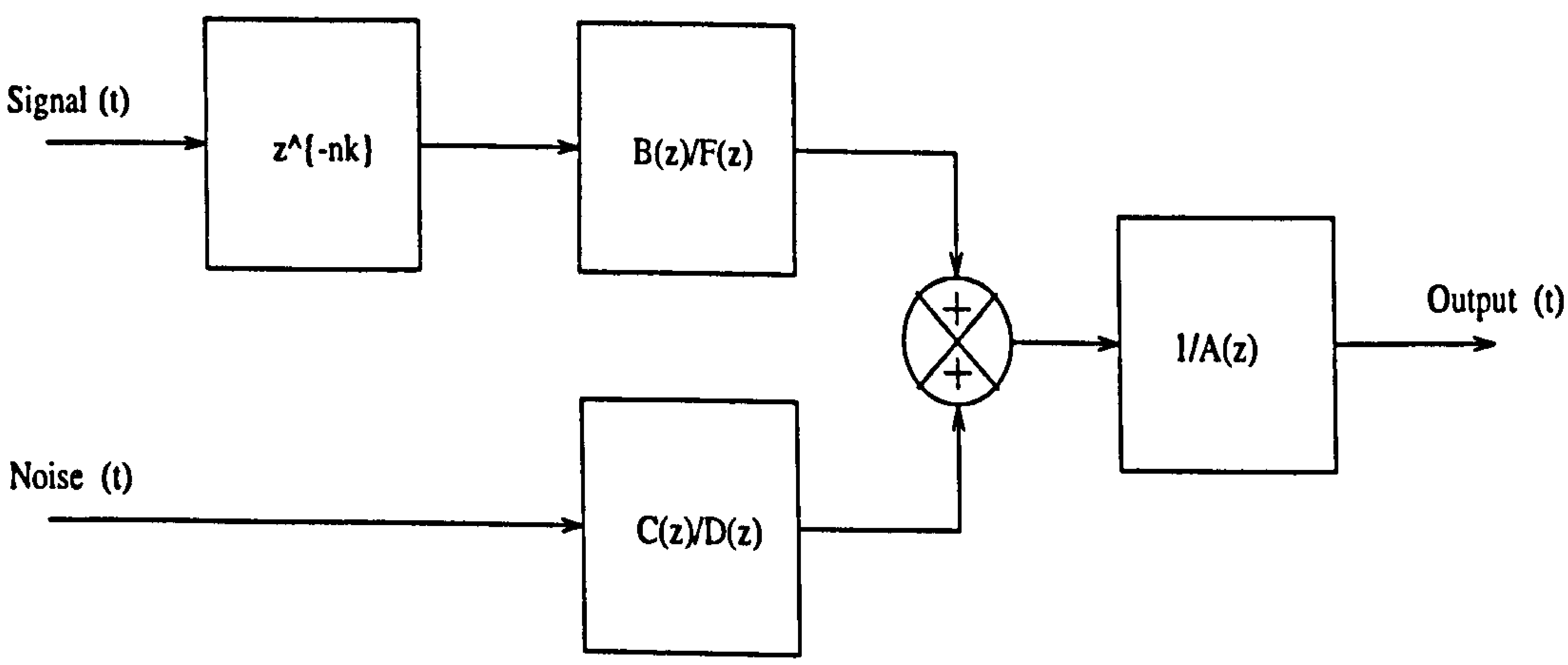
If the unit is used with a closed loop control system then in the ideal case the output will perfectly mimic the input regardless of changing conditions and even of non-linear effects occurring in the device. In practice there are limitations, for instance the gain of the system must be low enough to ensure stability under all circumstances which limits the "correcting power" available. Despite these limitations closed loop control techniques can lead to marked performance increases in some systems.

5.3.7 System Identification

System identification is the process of building a mathematical model of dynamic system based on observed data from the system. The observed characteristics of such a system are termed outputs and these are the result of stimuli. The stimulation to a system can be separated into two distinct classes, the user inputs, which allow external interaction, and disturbances. The latter can be further divided into observable disturbances and those whose effect can only be seen in the output of the signal.

After slight simplification, omitting the second class of disturbances, a generic system could would look like figure 5.10

Figure 5.10: Generic System Identification System



The system identification task is to identify the functions $A_{(z^{-1})}$,$B_{(z^{-1})}$,$C_{(z^{-1})}$,$D_{(z^{-1})}$ and $F_{(z^{-1})}$.

There are a number of simplifications to this general model that are traditionally given names, such as ARX where $C_{(z^{-1})} = D_{(z^{-1})} = F_{(z^{-1})} = 1$ and ARMAX where $D_{(z^{-1})} = F_{(z^{-1})} - 1$.

Matlab provides routines capable of carrying out the system identification process for the general problem, and thus by definition all the common structures. The identification process for a simple problem is described in more detail in chapter 2. Further detail on the actual process used by Matlab can be found in [34].

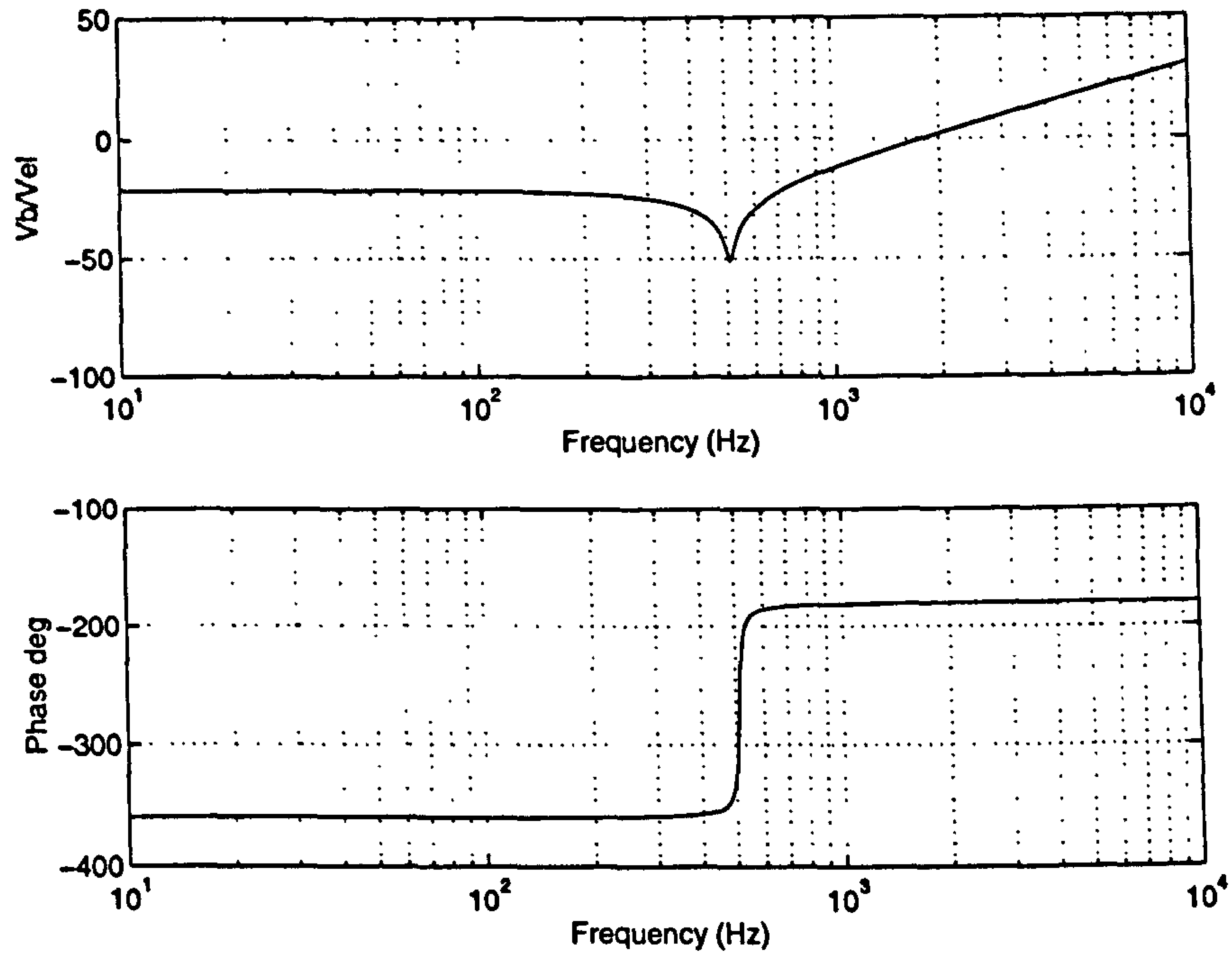
5.4 Conceptual Design of System

5.4.1 Concept in Detail

It is possible to produce a more linear loudspeaker by placing it in a feedback control loop. As discussed previously it is necessary to have some performance measure in order to make use of this technique, in this case a measure of the cone's movement. There are many different methods of observing the movement of the cone, some of which were considered briefly earlier. Of these methods perhaps the most desirable is to make use of the back EMF generated by the movement of the voice coil in the magnetic field.

There is a serious problem that needs to be overcome if this method is to be of any value in practice. If the full equivalent circuit is considered *figure 5.8* it is easy to see that simply including the speaker in a resistive bridge circuit will not produce the desired results. The voice coil's self inductance will produce a series resonance with the cone's mass (modelled by a capacitance). The result of this resonance is that the output of a resistive bridge will fall very dramatically at this resonance which occurs at an inconvenient point in the frequency band. The output expected from a resistive bridge is given in *figure 5.11*.

Figure 5.11: The Output of a Resistive Bridge Relative to the Cone Velocity



In addition to this the back EMF is related to the velocity by the factor BL. While L, the length of the windings, may be expected to remain constant within reason the value of B, the magnetic flux density, may not. The back EMF is given by Faraday's equation[13]:

$$N \oint E.dl = \frac{-d}{dt} \int \int B.dA \tag{5.14}$$

This can be simplified by assuming that the vectors align exactly but this is not strictly necessary. It is not hard to see that should B decrease then E will decrease accordingly.

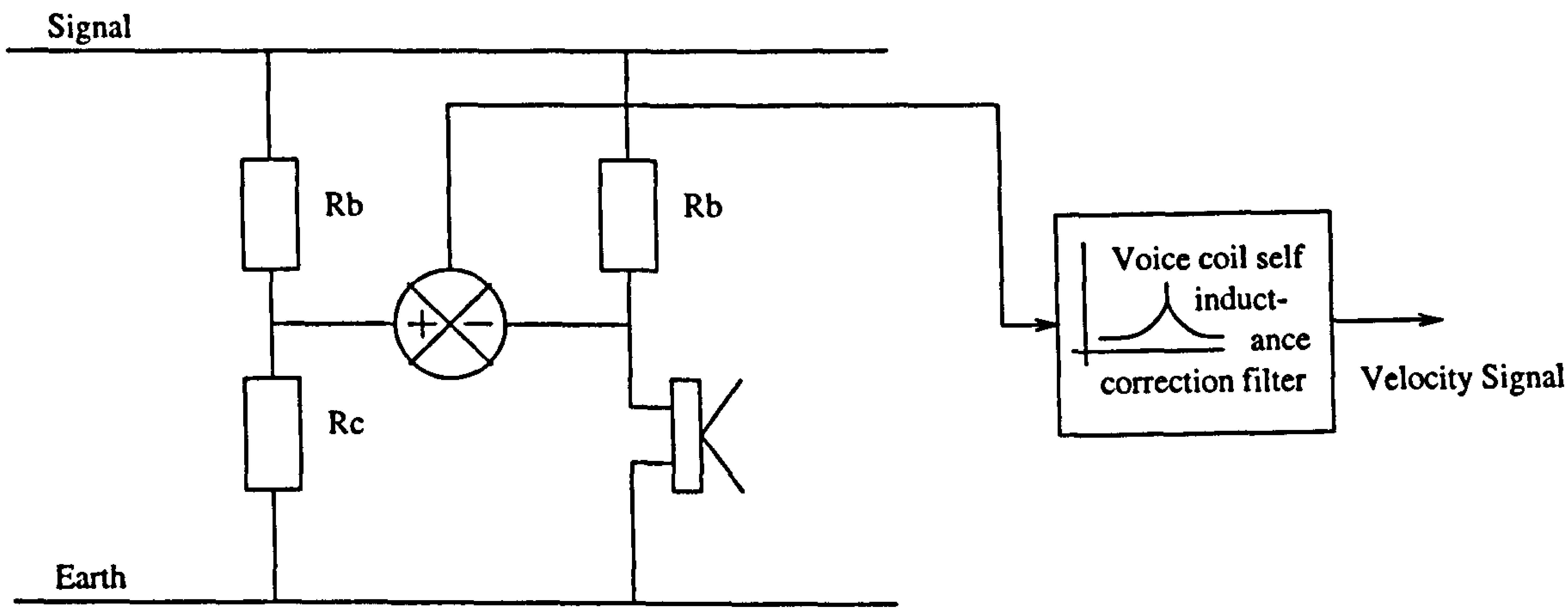
Unfortunately the back EMF generated in the coil of a speaker will reduce in proportion to B and thus the accuracy of the back EMF as a measure of velocity is impaired.

The temperature of the voice coil will change significantly during operation and thus a simple resistive bridge cannot be expected to remain in balance [6].

5.4.2 Compensating for the Voice Coil's Self Inductance

There are two options that may be considered to compensate for the voice coil's self inductance, one is to follow the bridge with a filter designed to compensate for the effects of the voice coil impedance, the other is to alter the bridge to cancel the effects. While following the bridge with a filter may appear the most natural route the high gains required and high "Q" factor mean that the filter will have to be very accurately defined if the output is to track the velocity with sufficient accuracy.

Figure 5.12: A Resistive Bridge with a Corrective Filter



As an alternative it seems that it should be possible to design a more complicated bridge arrangement that will overcome this problem - the details are considered a little later. In practice this method was chosen since it promised to be considerably easier to implement and potentially yield better results. With either solution there remain problems that need to be overcome.

- Tuning

  There will be several elements that will require tuning. If carried out manually this would be an expensive process.

- Drift

  The tuning process may be quite critical and it is essential that the compensation does not drift or the feedback control will have a very damaging effect on the sound quality.

- Speaker

  The physical constants governing the operation of the speaker, for example the voice coil resistance, may well change over time and with different operational environments

- in fact it is likely that this could prove more troublesome than the drift of the bridge components.

Fortunately there is a solution to these problems. If system identification techniques are used the information required about the loudspeaker can be obtained in real time and without direct human intervention by analysing the signal fed to and a suitable output of the loudspeaker. The system identification techniques are not dissimilar to the adaptive control that will be used in chapter 7 for acoustic compensation.

The proposed solution is therefore to make use of system identification techniques to establish the coefficients for the chosen solution required to compensate for the series resonance. The output of this can then be used to measure the movement of the cone and thus a correction signal can be fed to the input.

Obviously the system identification does not identify the parameters required for the solution but rather those of a model for the test system. Some processing will be required to make use of this information. If the equivalent circuit of the loudspeaker is used it is easy to establish an s-domain expression for the unit's impedance in terms of the bulk parameter components that make up the model. Observing the speaker using system identification techniques will typically produce a model for the real speaker in the z-domain, but this can be translated to a directly comparable form. There are well known methods that will allow the conversion of the system identification output into the s-domain where it will be almost equivalent to the model derived from the bulk parameters (within the limitations of the model and universally scaled by an arbitrary factor.) Using equivalence techniques a set of simultaneous equations will be obtained, the solution of which yields the values of the components in the loudspeaker equivalent circuit model. From this knowledge it is possible to design countermeasures for the the effects of the voice coil self inductance.

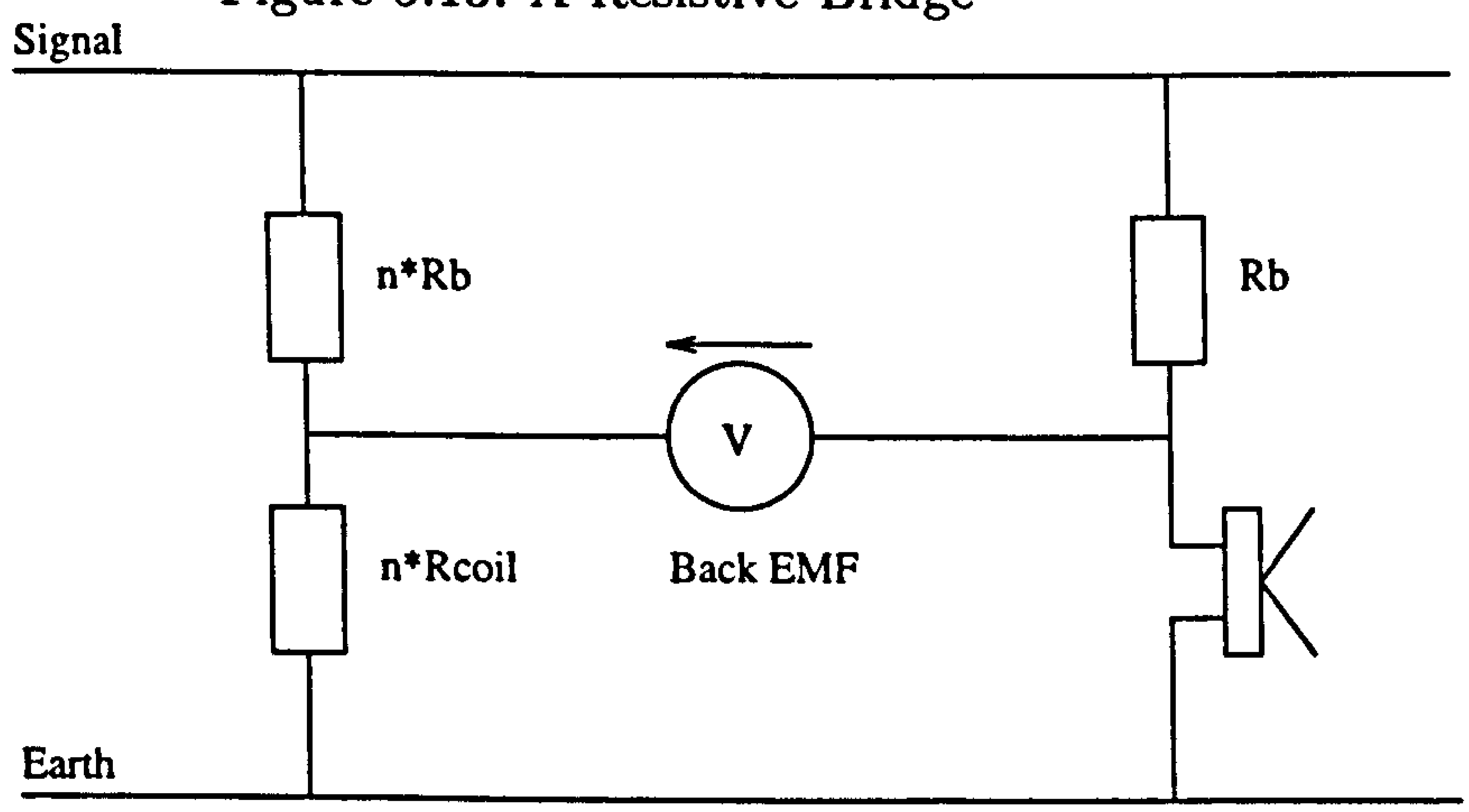
As an alternative, the results of the system identification process could be used to tune an open loop control system. For instance the value of the voice coil resistance measured using the system identification could be used to compensate for the effects of voice coil temperature fluctuations by adjusting the gain and thus yield a more consistent output level. This method is inherently limited when compared to closed loop techniques.

It is envisaged that most of the processing will occur in the digital domain; it is possible to make use of electrically variable components but these types tend to be particularly beset by parasitics. Digital simulations can be varied and are inherently perfect in their operation.

The Bridge Circuit

The intention is to produce a bridge circuit that will yield a signal proportional to the cone velocity. Considering the electrical analogy *figure 5.8* then the velocity is equivalent to the voltage across the capacitance which represents the cone mass. If the circuit is placed in a normal resistive bridge, such as that in the figure below:

Figure 5.13: A Resistive Bridge

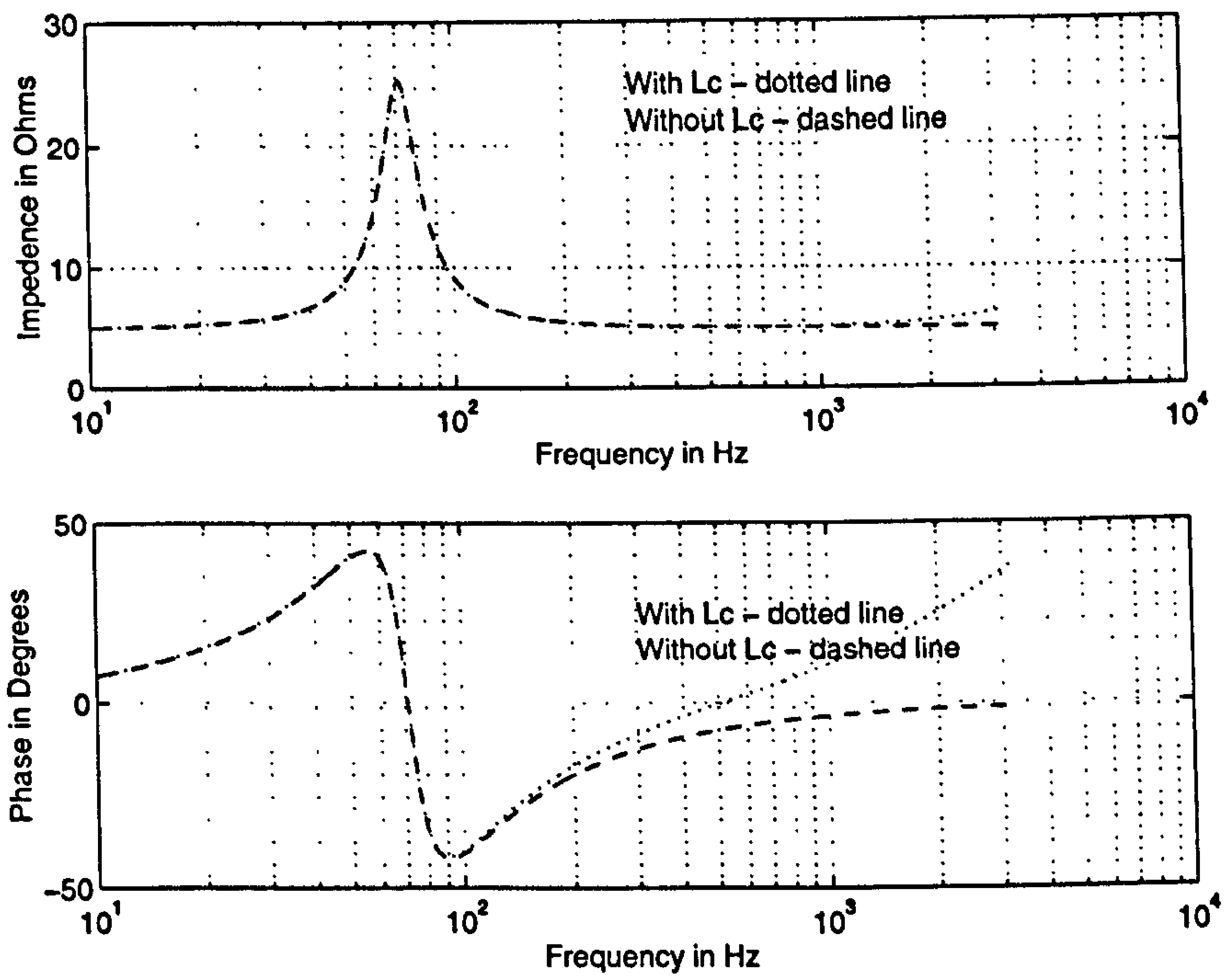


n.b. Rcoil is the DC resistance of the loudspeaker's voice coil.

It is can be seen that the bridge output would be related to the velocity apart from the effects of L_c. At low frequencies the reactance of the capacitance will be high and thus the effects of L_c, whose reactance is low at low frequencies, will be limited. As the driving frequency rises the roles are reversed and the effect of L_c becomes serious.

The effects on the driver impedance can be seen in *figure 5.14*:

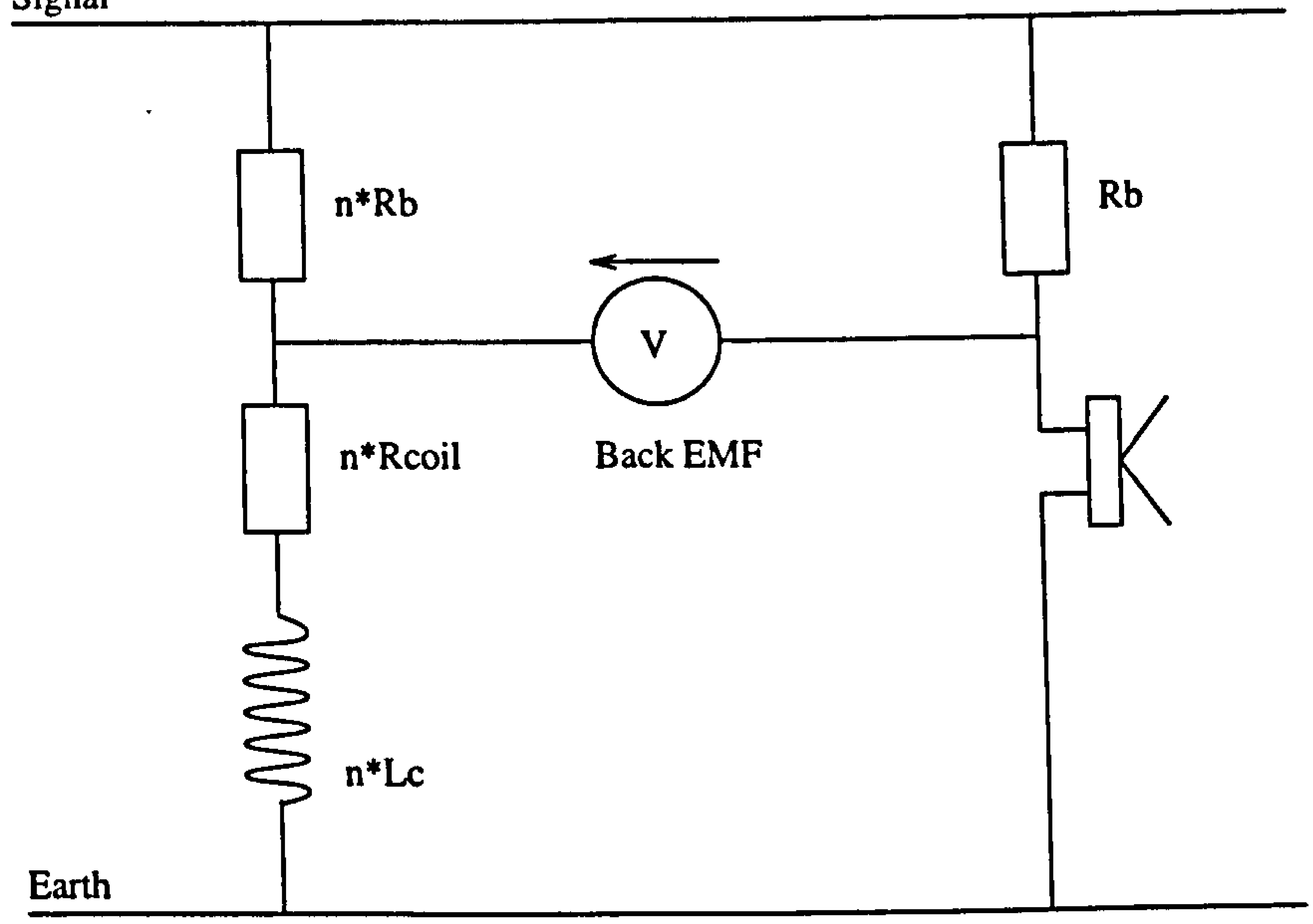
Figure 5.14: Impedance of a Driver With and Without Lc



Ignoring possible changes in temperature of the voice coil, the most serious problem occurs where the impedance of the driver becomes purely resistive and very nearly equal to the DC resistance of the voice coil. Under this condition the output of the bridge will be very low.

Fortunately it is possible to resolve this problem by using a more complicate bridge structure that is designed to cancel both the effects of the inductance and the resistance. All that is required is a mimic of the effects of the inductance at high frequencies on the balancing leg of the bridge and the desired output will be obtained. (At high frequencies the loudspeaker may be thought of as a resistance, inductance and voltage source in series.)

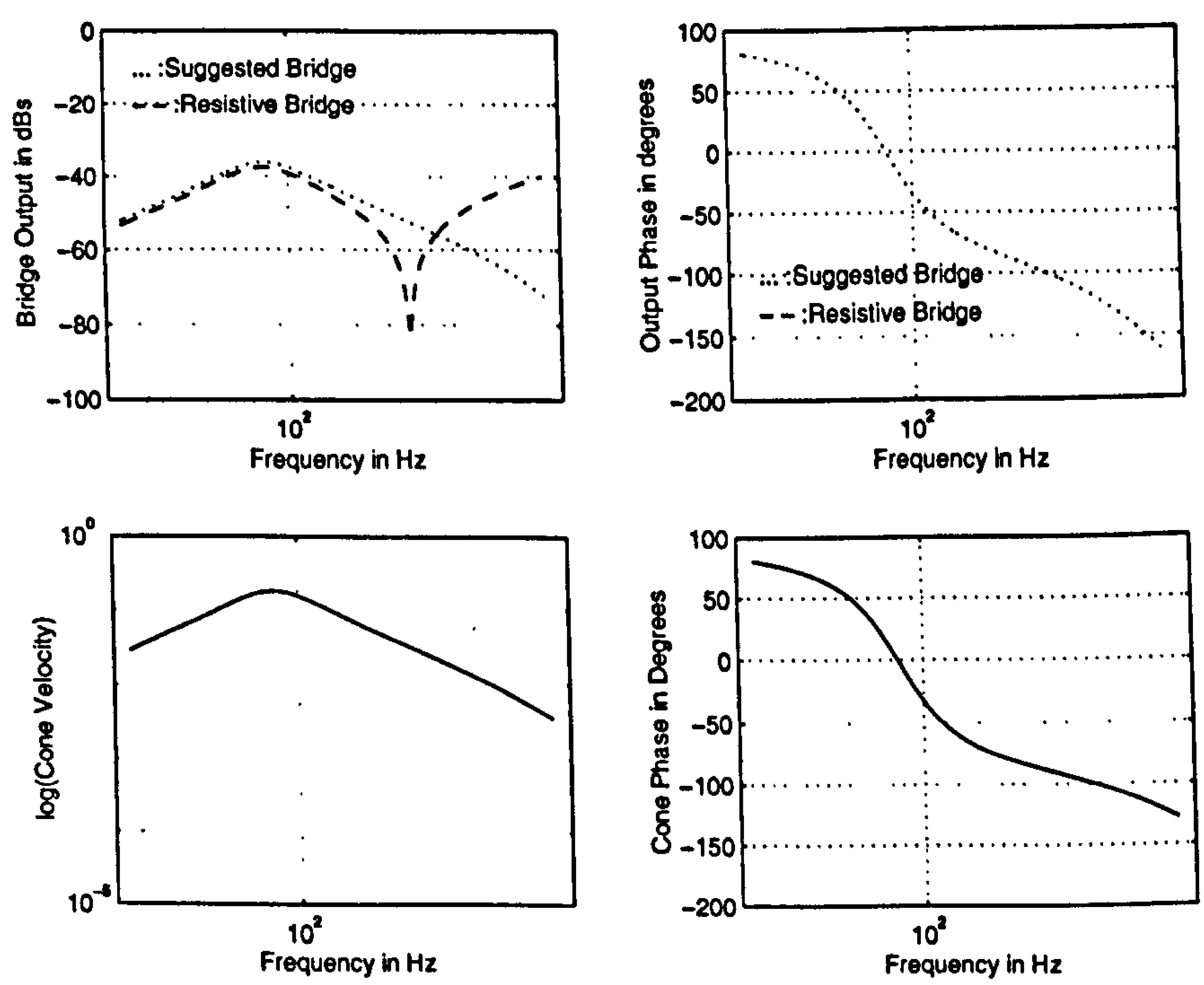
Figure 5.15: A Complex Bridge



n.b. Rcoil is the DC resistance of the loudspeaker's voice coil.

If the concept is simulated using Matlab it can be seen that the bridge output now accurately mimics the cone velocity. It should be noted that there will be parasitic effects that will mean the solution will not be quite as good as the simulation indicates. The values required to be obtained from system identification are L_c and R_c.

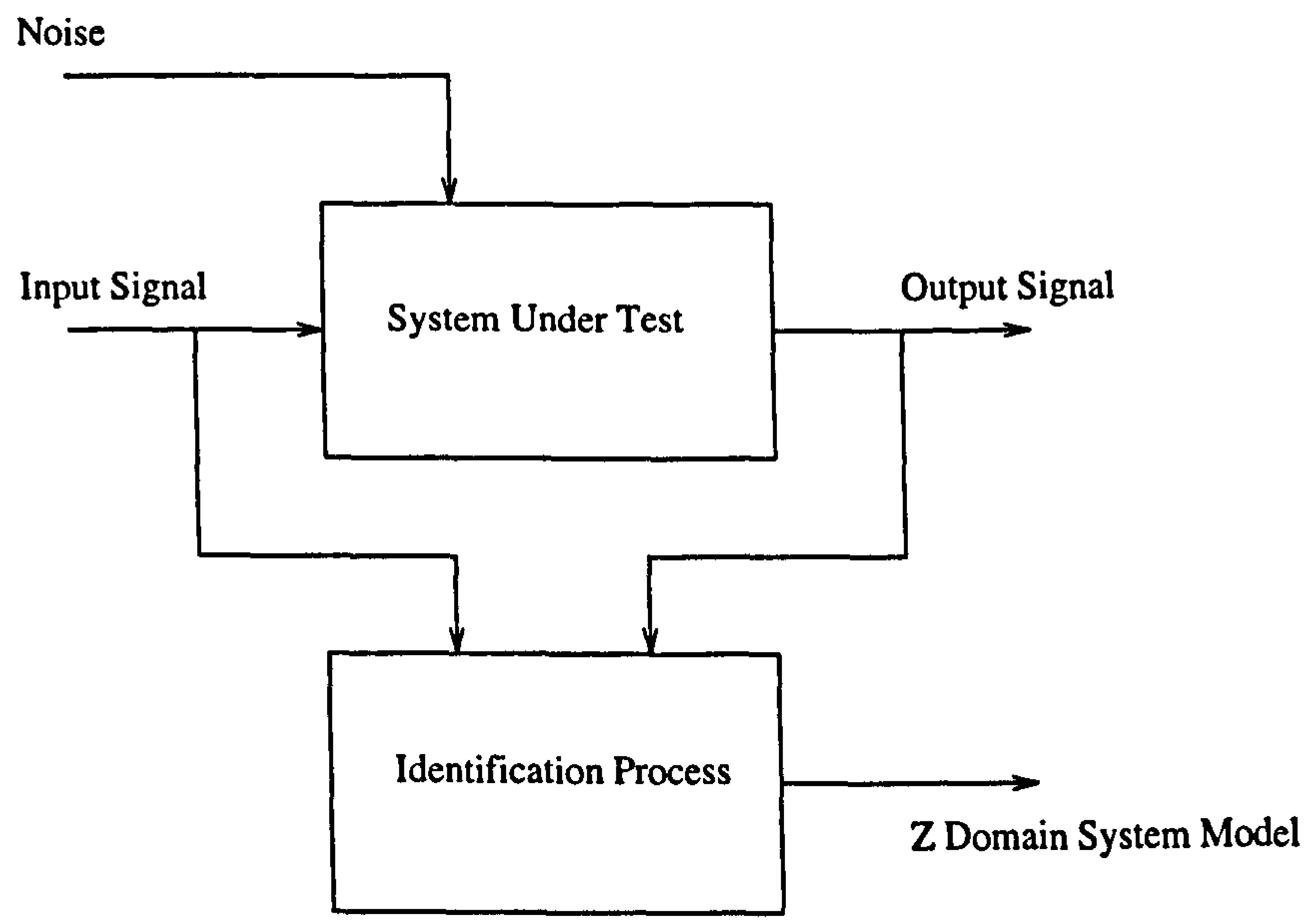
Figure 5.16: Voltage Output Relative to Cone Velocity for a Complex Bridge



5.4.3 Identification of Parameters in the Bridge Circuit

System identification is not a new concept, the ideas are well understood for common arrangements. If the system to be identified can be expressed as a black box with a single input and output then the process required to identify the system is well documented.

Figure 5.17: System Identification



As stated previously the model obtained by system identification techniques can be transformed so that it is in a form that is almost equivalent to the appropriate s-domain function obtained from the bulk parameter model. Since the two representations are almost equivalent (related by a consistent factor) a number of simultaneous equations can be formed which, when solved, will give values to the individual components in the bulk parameter model.

It is well known that solution of simultaneous equations is not necessarily trivial, especially when the equations concerned are not exclusively linear. It is therefore essential that the proposed solution yields a set of equations that is sufficiently easy to solve simultaneously that it can be carried out automatically in real time.

There are a number of factors that need to be considered in addition to the previous point: All of the systems will require some additional information in order to allow the identification to proceed, some may require more than others, and the expression of the identification problem may may give better resolution and accuracy with some parameters then others. Obviously any method favouring those parameters required is desirable.

The methods considered were, impedance identification, half bridge based identification and full bridge based identification. Looking at each of these in more detail:

- Impedance identification

  Summary of requirements:

  - Voltage Applied to Speaker
  - Current Through Speaker

  The requirements can be expressed, referring to figure 5.13, as the voltage drive to the speaker leg of the bridge and the voltage at the centre tap of the bridge. (The current can be calculated from the difference between there two voltages since the bridge resistance is assumed ideal. Note that this assumption may allow undesirable parasitic effects.)

  The electrical impedance is the simplest appropriate expression that can be formed to describe the speaker and thus the easiest to extract the parameters from. Initially this appeared a different type of problem and thus this solution was discarded. On later re-evaluation it seems likely that if the voltage were considered the "input" and the current the "output" the system could be identified by normal means. This method would offer excellent access to the speaker parameters. Since the alternative method initially favoured returned the results required this method was not re-analysed in detail.

- Half-bridge Identification

  Summary of Requirements:

  - Voltage Applied to Bridge
  - Voltage at the Centre of the Loudspeaker Leg

  One leg of a bridge is directly amenable to system identification methods. (That is a resistance connected in series with the loudspeaker. The connection between the speaker and the resistance is taken as the output.) This is a system with a normal input and output and a transfer function that is a lot less complicated than that of the entire bridge. This method was chosen since it was initially considered to be the simplest system directly amenable to system identification.

Considering only the loudspeaker leg of the resistive bridge shown in figure 5.13 and using the bulk parameter model for a loudspeaker, shown in figure 5.8 the transfer function for the half bridge system in terms of the bulk parameters in the "s" domain can be calculated:

$$\frac{V_{out}}{V_{in}} = \frac{k(a_3s^3 + a_2s^2 + a_1s + a_0)}{k(b_3s^3 + b_2s^2 + b_1s + b_0)} \tag{5.15}$$

where k is arbitrary and:

$$a_3 = L_cM \tag{5.16}$$

$$a_2 = R_cM + L_cB_c \tag{5.17}$$

$$a_1 = R_cB_c + L_cK + B^2L^2 \tag{5.18}$$

$$a_0 = R_cK \tag{5.19}$$

$$b_3 = L_cM \tag{5.20}$$

$$b_2 = M(R_b + R_c) + L_cB_c \tag{5.21}$$

$$b_1 = B_c(R_b + R_c) + L_cK + B^2L^2 \tag{5.22}$$

$$b_0 = K(R_b + R_c) \tag{5.23}$$

Where V_{out} is the voltage at the tap halfway down the leg of the bridge and V_{in} is the input voltage.

When split into simultaneous equations there will be seven equations, one being repeated, and there are eight variables. Thus it is possible that a non-unique solution can be obtained. It is no surprise that a single solution is unattainable given the information, even if the speaker were no more than a simple resistance its value could not be calculated without knowledge of R_b.

- Bridge Identification

  Summary of Requirements:

  - Voltage Applied to the Bridge
  - Voltage Developed Across a Resistive Bridge

  There is no need to include a full bridge in hardware, it is sufficient to have the loudspeaker side, the remainder may be implemented digitally. Implementing the bridge completely in hardware may produce inferior results due to the additional parasitic effects that will be experienced. In addition if the balance resistance is simulated digitally it can easily be altered to take into account changes in the voice coil resistance due to temperature effects.

A resistive bridge circuit also has a simple input and output but the analysis of the equivalent circuit yields a substantially more complicated transfer function. This means that the set of equations that would have to be used to determine the parameters from the output of the identification process are complicated. The order of the system to be identified is higher than for the impedance solution which will cause the identification process to take longer. The effects of the series resonance are particularly evident when analysing the output of a resistive bridge which might be taken as an indication that the identification would favour those parameters that are relevant.

$$\frac{V_b}{V_{in}} = \frac{k[a_3 s^3 + a_2 s^2 + a_1 s + a_0]}{k[b_3 s^3 + b_2 s^2 + b_1 s + b_0]} \tag{5.24}$$

where k is an arbitrary constant and:

$$a_3 = L_c M R_b \tag{5.25}$$

$$a_2 = (R_c M + L_c B_c) R_b - R_c R_b M \tag{5.26}$$

$$a_1 = (R_c B_c + L_c K + B^2 L^2) R_b - R_c R_b B_c \tag{5.27}$$

$$a_0 = R_b R_c K - R_c R_b K = 0 \tag{5.28}$$

$$b_3 = (R_c + R_b) L_c M \tag{5.29}$$

$$b_2 = R_c R_b M + R_b^2 M + (R_c + R_b)(R_c M + L_c B_c) \tag{5.30}$$

$$b_1 = (R_c + R_b)(R_c * B_c + L_c K + B^2 L^2 + R_b B_c) \tag{5.31}$$

$$b_0 = (R_c + R_b)(R_c K + R_b K) \tag{5.32}$$

Thus it can be seen that there are seven equations. If R_b^2 is taken to be a separate variable from R_b etc. and B^2L^2 is taken to be a single variable in order to effectively linearize the simultaneous equations then there are ten variables including k. It is essential that the simultaneous equations are linear if the general rules for the existence of a solution are to be applied. (The result of this simplification will be that the variable R_b^2 will not necessarily be exactly equal to the result of squaring the variable R_b, which is of no great significance. A more precise mathematical notation would be to make R_b^2 a completely different linear variable from R_b, but the current notation is considered sufficient.) Even if R_b (and therefore R_b^2) were given the equations will not be generally solvable.

The half bridge method was chosen because the adaption problem can be expressed in terms of a conventional system identification task but the problem of extracting the parameters remains tractable as against the full bridge method where the simultaneous equations are not generally solvable.

If the output of the identification process and analysis of the bulk parameter model are as follows:

$$\frac{V_{out}}{V_{in}} = \frac{k[a_3s^3 + a_2s^2 + a_1s + a_0]}{k[b_3s^3 + b_2s^2 + b_1s + b_0]} \tag{5.33}$$

$$\frac{V_{out}}{V_{in}} = \frac{([L_cM]s^3 + [R_cM + L_cB_c]s^2 + [R_cB_c + L_cK + B^2L^2]s + [R_cK])}{([L_cM]s^3 + [M(R_b + R_c) + L_cB_c]s^2 + [B_c(R_b + R_c) + L_cK + B^2L^2]sK(R_b + R_c)])} \tag{5.34}$$

Then the following simultaneous equations are obtained:

$$ka_3 = L_cK \tag{5.35}$$

$$ka_2 = R_cM + L_cB_c \tag{5.36}$$

$$ka_1 = R_cB_c + L_cK + B^2L^2 \tag{5.37}$$

$$ka_0 = R_cK \tag{5.38}$$

$$kb_3 = L_cM \tag{5.39}$$

$$kb_2 = M(R_b + R_c) + L_cB_c \tag{5.40}$$

$$kb_1 = B_c(R_b + R_c) + L_cK + B^2L^2 \tag{5.41}$$

$$kb_0 = K(R_b + R_c) \tag{5.42}$$

The solutions obtained were:

$$K = \frac{k}{R_b}(b_0 - a_0) \tag{5.43}$$

$$B_c = \frac{k}{R_b}(b_1 - a_1) \tag{5.44}$$

$$M = \frac{k}{R_b}(b_2 - a_2) \tag{5.45}$$

$$L_c = \frac{R_ba_3}{b_2 - a_2} \tag{5.46}$$

$$R_c = \frac{R_ba_0}{b_0 - a_0} \tag{5.47}$$

$$B^2L^2 = k\left[a_1 - \frac{a_0(b_1 - a_1)}{b_0 - a_0} - \frac{a_3(b_0 - a_0)}{a_2 - b_2}\right] \tag{5.48}$$

Initially this may appear inadequate but considering that only L_c and R_c are required then it is sufficient. It may be assumed that R_b will be known. In practice this assumption is not required, R_b may be set to any value provided that it is used consistently. (It does not have to be equal to the value of the real resistor in series with the loudspeaker.)

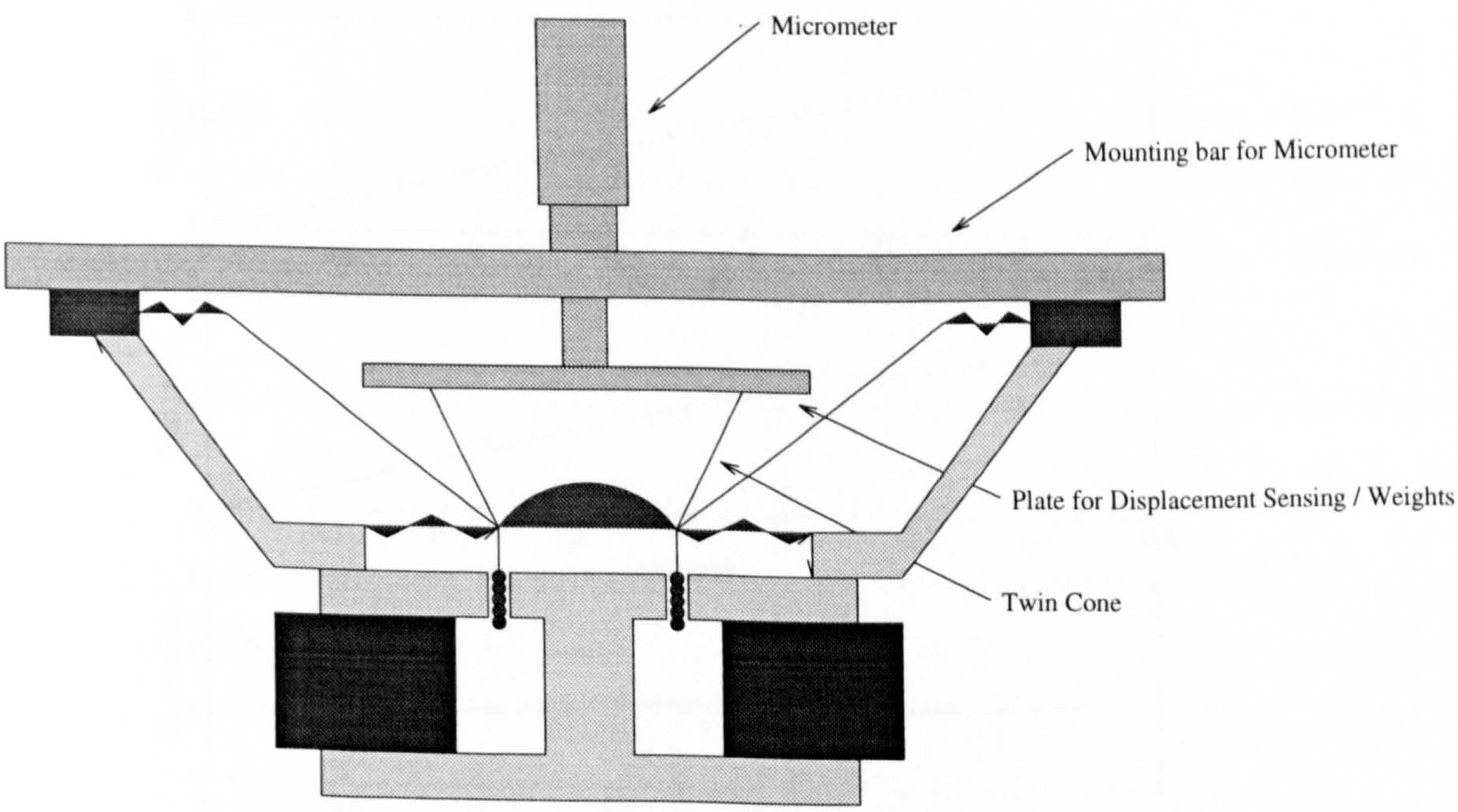
5.4.4 Variation in the Magnetic Flux Density

As stated previously the back EMF will fall with the value of B and thus the system's estimate of the cone velocity will be low. The value of B will fall either at the extremes of excursion, where the voice coil is outside the constant flux density area of the magnetic gap or with age as the magnet deteriorates. The first of these is the more serious problem since the effects will cause non-linear type distortions, the later occurs slowly [12]*pp 139*.

Realistic estimates of the changes in the magnetic flux density with the movement of the coil proved difficult and thus a practical investigation was attempted in order to the variation that might be expected. The speaker used for this experiment was the 8 inch general purpose device sourced from RS. employed throughout the project. It should be noted that this is not a HiFi grade device and it is expected that high quality devices would demonstrate a much more even magnetic flux density over the usable excursion.

Initially the driver was held, cone uppermost and a bar fixed across the front of the chassis. Into the bar a micrometer was fixed to measure the cone displacement. The twin cone provided a convenient flat surface to rest a plate on to take measurements against. A simple multi-meter was used in resistance mode to determine if the micrometer touched the metal disk on the cone.

Figure 5.18: Test Equipment



First, various masses were placed on the plate and the deflection measured using the micrometer. The resulting measurements allow the force at specific deflections to be calculated. Next the masses were removed and the same deflection was achieved passing a DC current through the voice coil. Any deflections could have been used but using the same deflection for both measurements removes the need for interpolation and the associated errors.

When plotted the applied mass against deflection demonstrates a slight deviation from the ideal straight line, with the spring constant increasing with deflection as would be expected. This is a non-linearity that a feedback control system might hope to reduce.

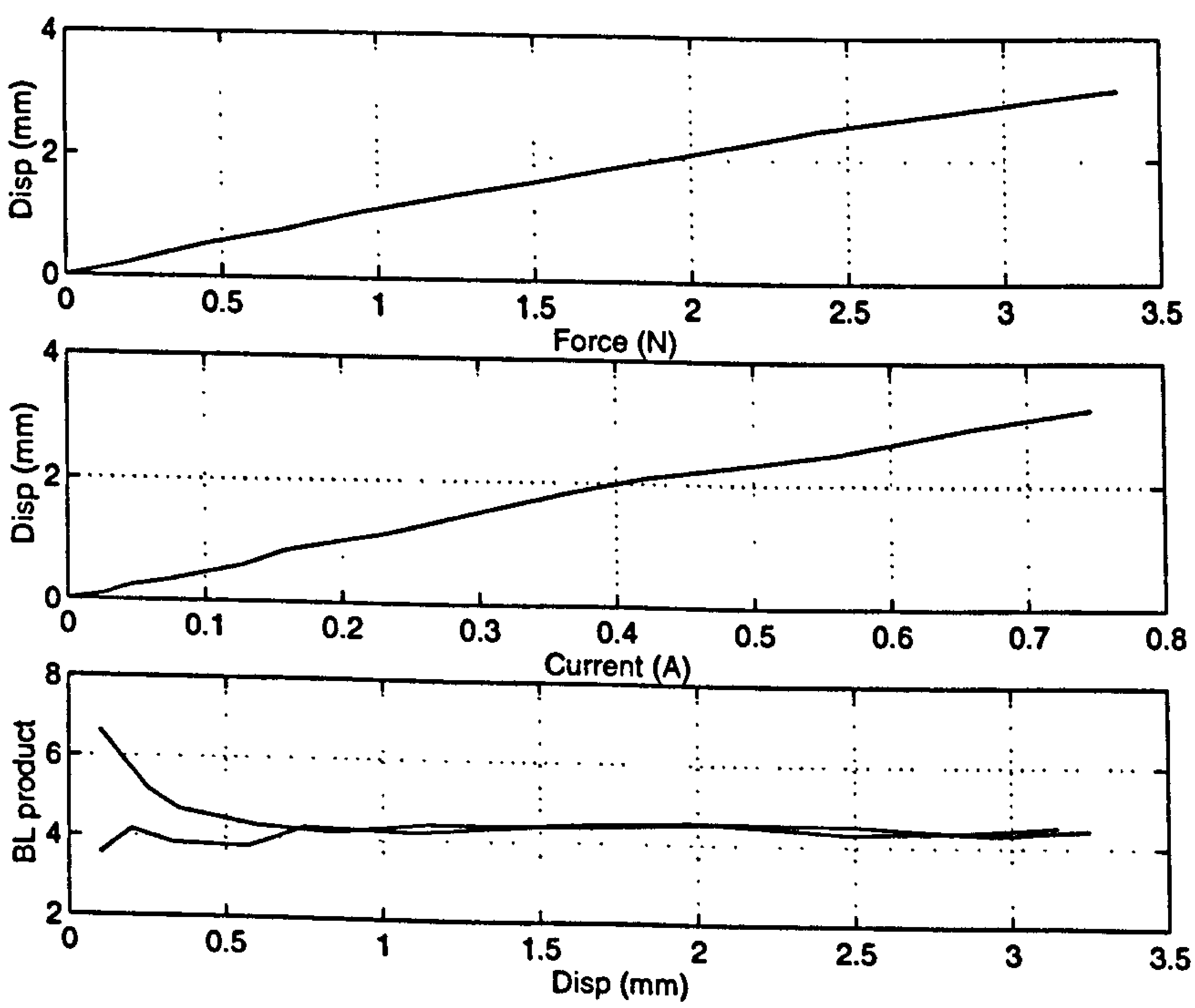
Since the displacements for the two sets of results are identical it is trivial to produce a plot of the BL product against displacement using:

$$BL = \frac{f}{i} \tag{5.49}$$

Where f is the force in Newtons and i is the voice coil current in Amperes.

In practice the results are limited due to the crude methods, two complete runs of the experiment were carried out to establish some indication of the error to be expected. On the first run the mass was added from the least to the greatest and on the second a substantial mass was added and removed before the start, which may go some way towards explaining the differences in the results. Many of the errors of measurement are particularly significant at low displacements and therefore it may well be better to discard these results. The measurements obtained are in the appendices, section 9.7 and are plotted in *figure 5.19.*

Figure 5.19: Linearity of the BL Product with Cone Excursion



The conclusion that must be drawn is that, within the error limits of the experiment there is no indication of change of the BL product over the displacements measured. (The maximum displacement of 3mm from rest is substantial for this size of speaker.) As a rough indication the error of the experiment is of the order of 5% based on the maximum difference recorded between the two runs at a significant displacement.

A more detailed estimate of errors is possible if an estimate is made of the error attached to each measurement is made. The masses used, particularly the values below 200g were very accurate and thus unlikely to contribute to the error significantly. The accuracy of the current measurement is probably of the order of ±1mA; given that the mid-range currents were of the order of 100mA this represents an error of 1%. The location measurement has an error of about ±0.02mm; given that the displacement at mid range is of the order of 1.5mm this represents an error of about 1.5%.

A cumulative estimate of the error in the displacement/applied mass measurement is about 1.5% and of the displacement/current measurement is 2.5%. Thus this method yields an estimate of error for the final results set of about 4%. The errors at low excursions will be much greater.

5.4.5 Effects of Changing Voice Coil Resistance

HiFi loudspeakers are inefficient devices, it is unlikely that more that 1% of the input power will be converted into sound with the majority of the remainder being dissipated as heat in the voice coil. This results in the voice coil undergoing substantial changes in temperature since it is difficult to dissipate heat effectively from the voice coil.

According to Birt [6] voice coils of modern loudspeakers can have maximum operating temperatures in excess of 200°C. Given that the copper typically used to wind these coils has a temperature coefficient of 0.4% then the voice coil resistance can be expected to change by a factor of close to 2:1 over its operating temperature range. Birt also states that the thermal time constant for the voice coil should be expected to be in the range of 40 seconds for a large LF driver to 1 second for an HF driver.

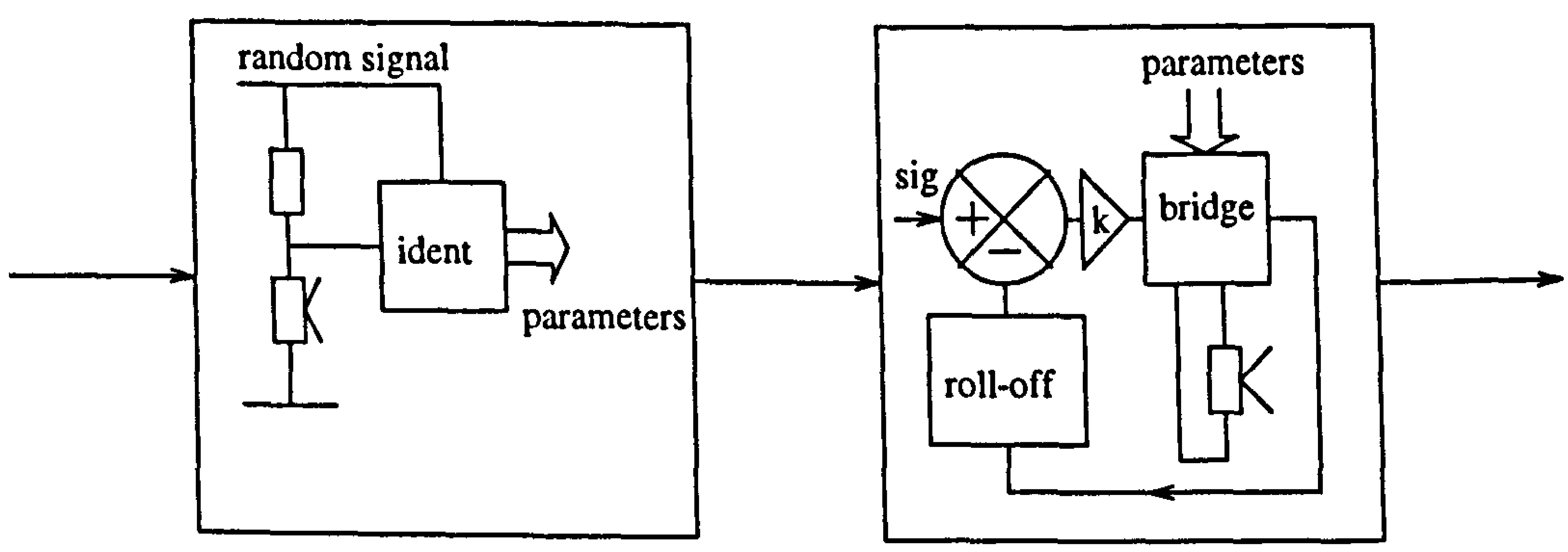
If a loudspeaker were placed in a bridge with a fixed resistive element balanced against the coil resistance, as suggested previously in order to determine the back EMF, then changes in the coil resistance of this magnitude will have a serious impact on the output. Fortunately, the system identification method proposed is able to identify the value of R_c. Provided that the identification of R_c can be carried out sufficiently quickly to ensure that the value is a good representation of the coil resistance then the changes in voice coil temperature should not represent a problem. Given that the thermal time constant of a large LF driver is of the order of 40 seconds the identification needs to take place in much less than about 20 seconds for a more average device.

In conclusion, given that the voice coil temperature time constant is of the order of 20 seconds for the LF loudspeaker, provided that the system identification process takes less than about 2 seconds then the errors should not be significant. The mechanical time constants of the loudspeaker are much less than 2 seconds which means that this target is realistic. The variation in output reported by Birt due to the changes in voice coil resistance will not occur in a system that makes used of closed loop control techniques.

5.4.6 Demonstration of Closed Loop Control Using the Back EMF

The concept can be demonstrated using the Matlab system identification and control toolboxes. The process has been demonstrated from the initial acquisition of the system parameters to an analysis of the closed loop system that is produced. In a real system there would have to be some additional control that decided when the parameters were sufficiently mature to allow the closed loop operation to begin.

Figure 5.20: Demonstration System



The demonstration was carried out in two very distinct stages, the first being the parameter estimation and the latter using that information in an evaluation of the simulated closed loop control system. In the parameter estimation stage a system identification was carried out on the loudspeaker using a half bridge type set up. The parameters were then translated for use in the closed loop control part of the system. (These processes are represented in the left-land box in figure 5.20.)

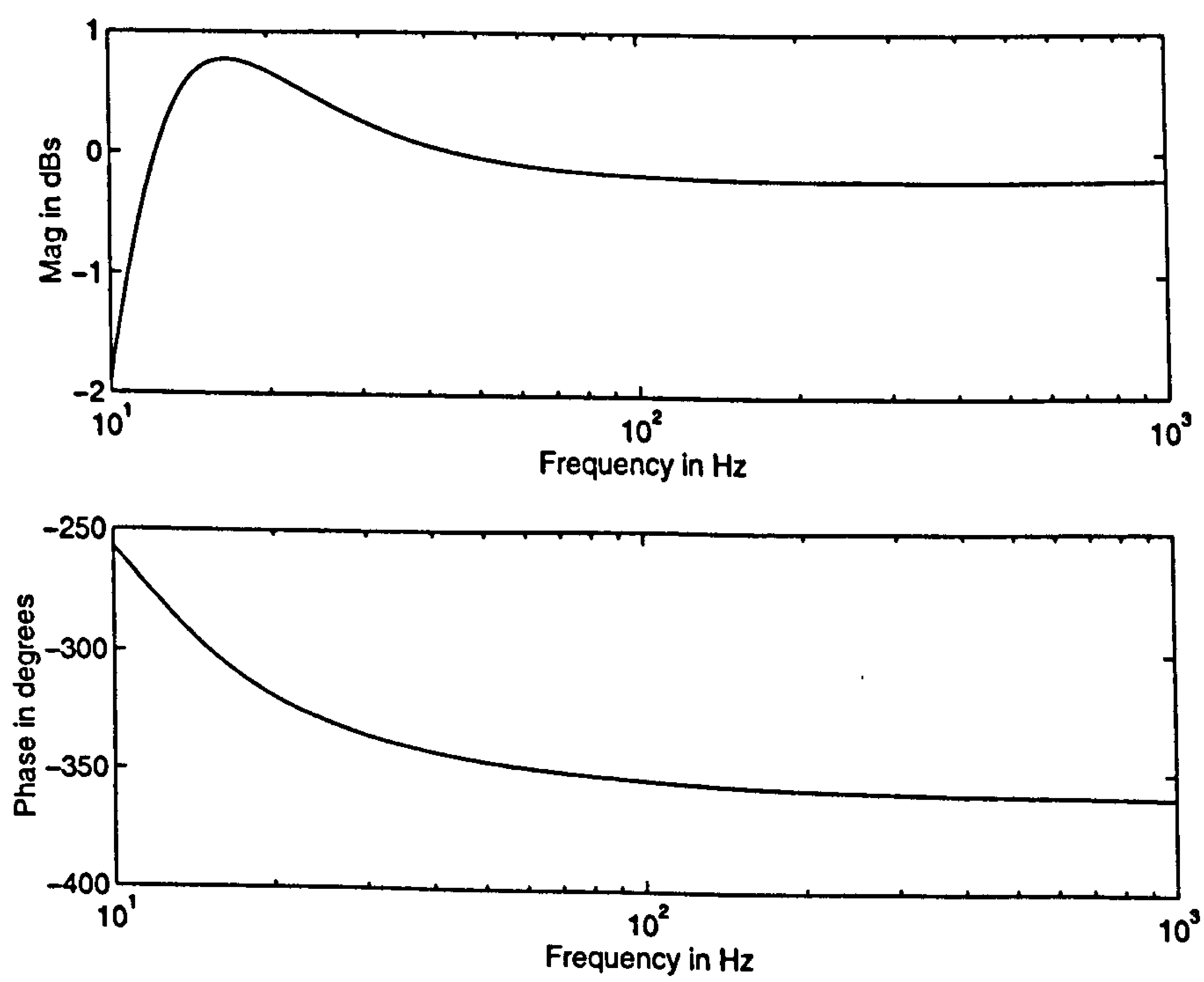
The parameters from the system identifier were used to set up a closed loop control system, as described earlier, around the loudspeaker.

The parameters obtained from the system identification stage proved accurate, which led to the good final results. (The error was much less than 1%. It is unlikely that such accuracy could be attained in a real system.) The feedback control loop gain was set to 5 as a trade off between stability margins and system response. It should be noted here that the output is taken to be the cone acceleration, which it has been shown earlier is directly proportional to the acoustic pressure produced.

In the demonstration the system identification process operated on 1000 samples at a sampling frequency of 44.5kHz (representing approximately 0.02 of a second of signal at the normal CD sampling frequency.) and thus is well inside the target maximum of 2 seconds. Even given that it may not be possible to make use of as efficient methods of system identification in real time as were used in the demonstration the target should be achievable.

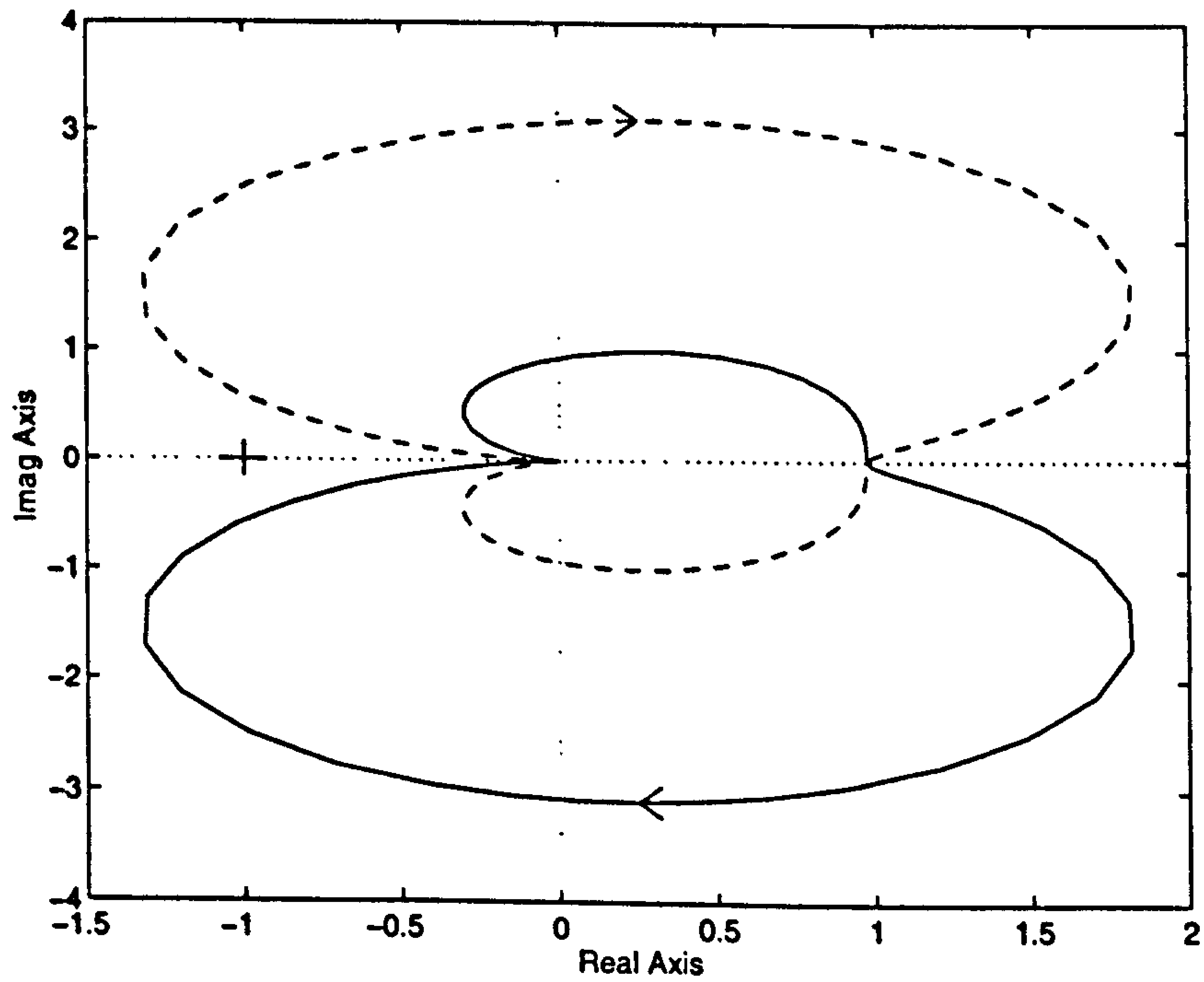
In practice the process was only carried out under simulation and thus the parameters were extracted from a model of the loudspeaker and likewise the control loop operated only on a model. Implementing the system using a real loudspeaker would have required DSP hardware in order for the process to operate sufficiently quickly.

Figure 5.21: Response of Simulated System



The stability margins are significant since the simulation will differ from the real system, particularly at high frequencies. If a Nyquist plot is obtained for a gain value of 5 then the results are promising. In practice the ideal values of the gain would very much depend on the operations of the real system. looking at the Nyquist plot, figure 5.22, the -1 point is well clear of the plot and thus stability should not represent a problem. (The system as simulated will not actually go unstable, regardless of the value of the loop gain. A real system would be affected by parasitic effects and thus would be more likely to suffer from instability if the loop gain were too high.)

Figure 5.22: Nyquist Plot for Simulated System



5.5 Considerations for Full Implementation

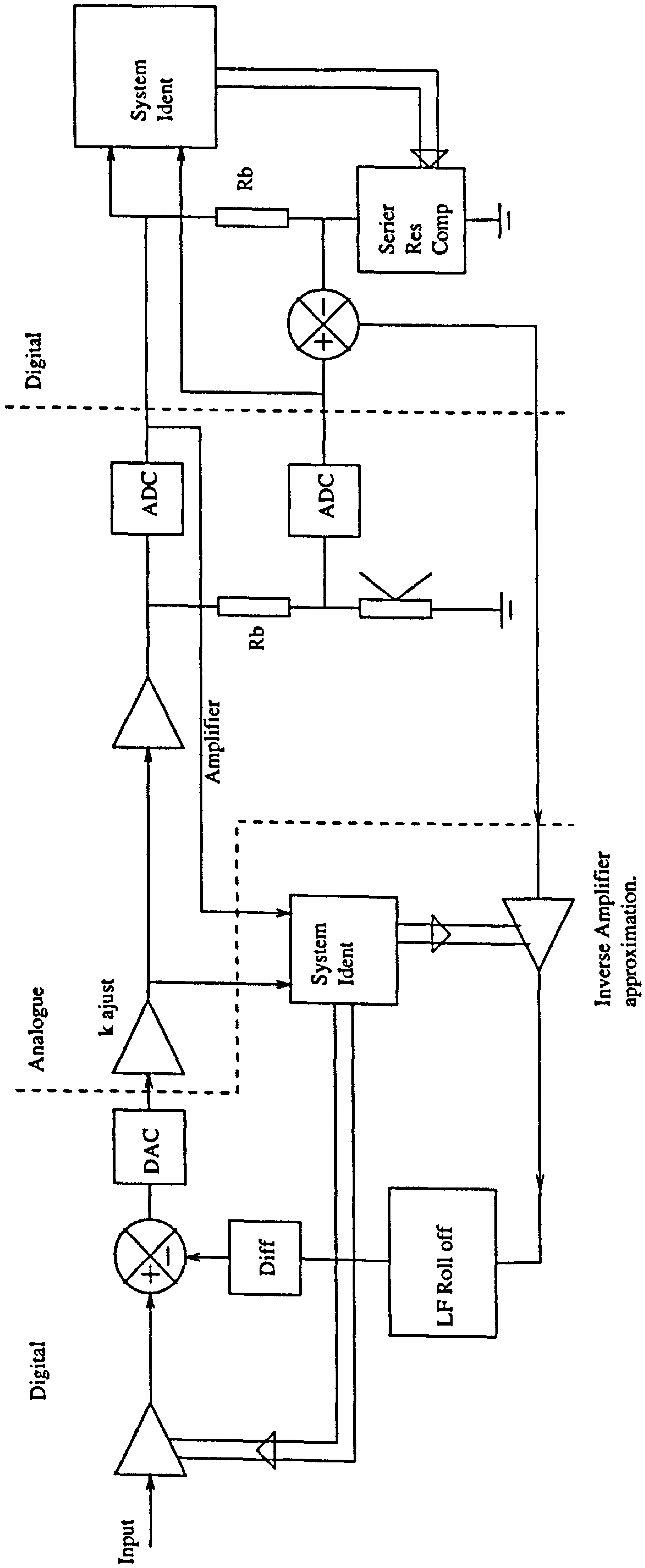
There are additional problems that need to be dealt with when the real implementation is produced. If an existing commercial amplifier were to be used which was separate from the back EMF speaker control device then it would have a variable gain that must be under the control of the user (the volume control), yet it must also be within the feedback control loop. (If the signal input is to be obtained direct from the source and subtraction to be done digitally) Thus, if no action is taken, the volume control will affect the stability of the system but not the volume of the output! There are two solutions to this problem, either the target signal will have to be obtained via an ADC from the output of the amplifier and the subtraction be performed at this high power stage or the direct signal could be scaled using the output signal of the amplifier as a guide. The former is highly undesirable. The latter option will still require an ADC to translate the output of the amplifier. Compensation will be required to ensure that the loop gain remains constant.Tone controls present a more difficult problem but audiophile grade equipment is rarely fitted with them. It would be possible to use system identification techniques to analyse the effects of the tone controls and apply a similar filter to the direct signal but the additional complication will probably reduce the reproduction quality.

A block diagram of this type of system excluding the supervisory control is given in *figure 5.23.*

Commercially the only realistic option is to combine the loudspeaker control function with an amplifier. This would allow access to the signal at various stages which would make the implementation less troublesome.

If the system is to be implemented largely digitally, as would appear desirable, then the volume control function will have to be handled digitally. The easiest control would be to have the level stepped by powers of two, since dividing by powers of two is convenient with a binary representation. This may be sufficient since some high grade amplifiers feature analogue gain controls with a relatively small number of discrete settings. The problem is that as the volume control level is decreased then information will be lost. (This is different from using an analogue attenuator since in the ideal world the analogue device will not lose any information, only scale the result. In practice, where the signal falls below the noise floor the information is lost.) Realistically, for digital systems, it must be considered permissible to discard information provided that the signal would have been below either the system noise threshold or the threshold of hearing. The use of 18, or better still 20bit DACs reduces the potential severity of the problem since information is not lost until the signal is being reproduced at a substantially lower level. The LSB of 18 and better yet 20 bit DACs are at a very low levels, comparable to the realistic noise floor of current amplification equipment. Processing would typically be carried out at a minimum of 24bits though modern DSP chips tend to operate with larger data widths and often use a floating point representation which removes a lot of the need for careful implementation of the DSP software which means that computational accuracy is unlikely to be a problem.

Figure 5.23: Full System using Commercial Amplifier



Further gains could be made if the speakers contained the analogue elements of the amplifier where the signal was merely processed digitally by a "pre-amplifier" and then sent digitally to a DAC located in the speaker.

It will be necessary for the supervisory control to update the value of R_c in the digital half of the bridge frequently to take into account the variation of the coil resistance with operating temperature. The changing coil resistance will also affect the effective open-loop gain of the system, but this can be compensated for digitally.

The full system excluding the supervisory control is given in the *figure 5.24*:

There are a number of additional problems and limitations that need to be considered:

- Ported Enclosures

  Many loudspeaker system currently on the market are of a ported type since this design can produce a better bass response from a small enclosure. Ported enclosures have different cone velocity/SPL profiles as it is not only the front surface of the cone that is radiating. It may be possible to compensate for this effect if the control system is set to roll off at the resonant frequency of the loudspeaker. There is little point in attempting to correct the response beyond this point since it is initially determined by the design of the enclosure, which is likely to be well behaved, and then falls very rapidly. In addition the bulk parameter model is more complicated to account for the effects of the port which will make the parameter extraction process more difficult or even impossible.

- Noise

  The additional complication of feedback control will inherently introduce additional noise. Since digital media are now the standard the SNR requirement has become very stringent. While it would probably not be possible to achieve comparable SNR figures with conventional amplifiers appropriate design will reduce this effect.

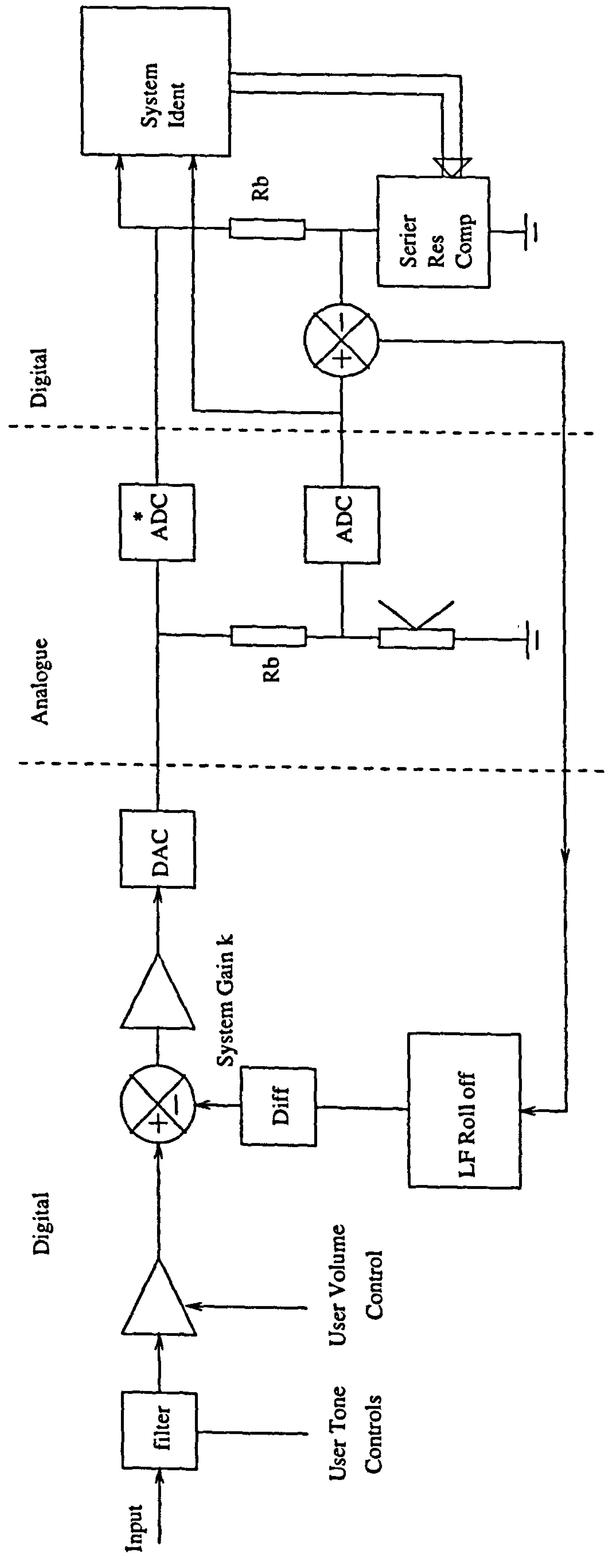
- BL value

  The cone velocity estimate is directly related to the BL value and thus if this value were to fall then the control system would "see" a low velocity and attempt to increase it to bring it's perception of the output into line with the desired input. For high gains the fact that the force on the cone will also decrease will the value of BL will have no effect on the output. At lower gains the effects may begin to cancel but on the whole the SPL will be expected to rise as BL falls rather than the reverse as is normal.

- Overload Handling

  Closed loop control can cause severe overloads in systems of this type if the limits are not set appropriately. Consider a powerful amplifier driving a smaller speaker and a signal that requires the cone to travel further than its maximum excursion. When the signal is applied the cone travels to its mechanical limit and is stopped, resulting in a signal from the sensor that indicates that the cone is not tracking its required movement. Correction is applied at a level that rapidly rises as the discrepancy increases, which can result in a significant overload condition developing in a situation that may not represent a problem in a more normal environment.

Figure 5.24: Full System, Integrated Amplifier



* Note this ADC not essential as the system gain is known and fixed but its addition will allow Rb to be sited close to the speaker which may be necessary.

5.6 Discussion of Results

The system for controlling the pressure output of a loudspeaker by making use of the Back EMF generated in the voice coil appears to be plausible. A simulated demonstration produced very good results but this inherently does not suffer from parasitic or noise effects. It is difficult to make any reasonable appraisal of the impact of these effects without implementing a real system.

Due to the constraints discussed then it is likely that the only way the system could be produced commercially would be to market it as part of an amplifier. Further performance gains could be obtained if the amplifier were integrated with the speakers but this is unlikely to be acceptable at this stage to the HiFi market in general.

At present the concept is only designed to work with sealed box type loudspeakers.

The next stage in the development of this type of system will be to produce a real prototype. There should be no difficulty reproducing the algorithms used from Matlab since these are well documented.

It is unfortunate that direct performance comparisons between the fixed filters currently used by some manufacturers and this concept are not possible. It is possible that the fixed filters will yield better results provided that the operation conditions are similar to those where the setting were calculated. As the conditions diverge further and the drivers age then potentially the dividends of feedback control will become fully apparent.

Chapter 6

Correction of Doppler Distortions due to Cone Motion

6.1 Introduction

A Loudspeaker produces sound by translating an electrical signal into a force that is used to move its cone and thus the surrounding air. An ideal Loudspeaker would be perfectly linear, in this case the SPL would relate directly to the voltage of the signal. In practice there are a great many ways in which real speakers fail to reach this target, some are linear effects and therefore relatively easy to correct, others are not. It is sometimes possible to design filters to correct even for non-linear distortions, depending on the exact nature of the problem. Loudspeaker designers go to great lengths to minimize the effects of non-linear distortion by looking carefully at the magnet system, the cone suspension and the cone breakup modes.

There is one non-linearity that even a "perfect" moving coil loudspeaker would suffer from since it results from the nature of the cone/air interface. Loudspeakers are often called on to reproduce many sounds of different frequency simultaneously, for instance that of a drum and a voice. Even an ideal driver can not reproduce such a composite signal without distortion due to the "Doppler" effect. The distortion will be apparent as the low frequency components modulating the high frequency components. Transmission to the listener imposes very little further distortion, since air may be considered linear, except at sound levels which would be painful.

The use of multiple drivers reduces the effects of the Doppler distortion since only the signals reproduced on the same cone will effect each-other in this way.

The Doppler effect is completely non-linear in that it generates frequencies not present in the original signal and thus it is difficult to design a filter that can compensate for its effects.

In this chapter a correction system is described. It appears that it will be possible to produce a real improvement in the level of this type of distortion.

6.2 Literature Survey

During the literature survey stage no result was found using the Bids database for the combinations of "Doppler" and "Loudspeaker" or "Doppler" and "speaker" but a reference to a methods for measuring the effects was found in [12]*pp 279* which in turn referred to a paper [3].

At a later date, when the literature survey was repeated, two new papers were found. In [60] Klippel describes a non-linear adaptive filter to reduce distortion caused by displacement sensitive parameters. He claims that this filter would also be effective at reducing the distortion caused by the Doppler effect. In practice this is not likely, all of the other displacement related distortions are of amplitude modulation type whereas distortion due to the Doppler effect is a frequency modulation type effect driven by the velocity. A second paper [66] is more interesting since this concentrates on the Doppler effect, which the author refers to as the effect of moving boundary conditions, and the effect of nonlinear acoustic wave propagation. At the conclusion of the paper Zoltogorski presents a block diagram for a filter to correct for Doppler distortion effect but takes the subject no further. This filter superficially appears quite different to that introduced later in this chapter.

6.3 Subject Revision

6.3.1 The Doppler process and How it applies to Loudspeakers

The Doppler effect can be most easily observed when an emergency services vehicle passes at speed, the apparent frequency of the siren reduces markedly. The same effect occurs in all emission systems and is responsible for instance for the red shift [1] observed in the emissions of nebula moving away from us. Since the speed of sound is relatively low the effect is observable even with "normal" systems.

The Doppler effect can be expressed as in [63]*pp 42*:

$$f' = \frac{(c-v)f}{(c-u)} \tag{6.1}$$

where

f	Emission frequency in Hz
f'	Perceived frequency in Hz
c	Speed of sound in m/s
u	Velocity of the source relative to the medium in m/s
v	Velocity of the listener relative to the medium in m/s

When reproducing a signal the cone of a loudspeaker is moving. If it is reproducing two or more signals simultaneously its movement will be the result of the superposition of the two resultant cone velocities, provided that the speaker is a "perfect" device. Taking the case of two signals; the cone can be regarded as a moving object, the result of the first signal, which is emitting the remaining signal and vice versa. Since every signal component results in a cone velocity then each signal component will cause “Doppler distortion” to every other.

The peak velocity of a loudspeaker cone is relatively small and thus the effect is not strong. The velocity falls above the resonant frequency of the driver/enclosure combination, this frequency is typically of the order of 50 - 100Hz. Since the cone velocity is much higher at low frequencies then the apparent distortion will probably take the form of the lower frequency component distortion the higher.

The Doppler distortion effect in speakers is well known but no attempt has been made previously to correct for it since the effects are small and the solution would be very difficult without the aid of modern DSP technology. The problem is, however, becoming increasingly acute with the current trend of decreasing size of the drivers causing the peak cone velocity to increase.

Doppler type distortion is a type of FM modulation whereas many of the other non-linear distortion mechanisms present in a loudspeaker result in AM modulation. For instance the result of changing BL product or suspension compliance with excursion will both result in an AM type distortion. One strategy to reduce AM type distortions is to use feedback control as discussed in the previous chapter, which will also obviously also combat linear effects

An interesting difference between FM and AM type distortions is that the signal will affect itself; for instance if the rate the cone centralizing force increases with excursion is not constant even a single sine wave signal is “compressed” (typically) as the excursion increases. For FM this self-distortion does not occur since the velocity associated with a component is an intrinsic part of the signal. In practice the LF components tend to cause the greatest excursion and velocity and thus these cause the highest levels of distortion to the remainder of the signal spectrum.

6.3.2 The Minimum Detectable Interval

In order to evaluate if Doppler distortion is audible a good knowledge of the minimum detectable intervals is required. The minimum detectable interval is the smallest change in pitch (frequency) that can be detected by a normal human. A table is reproduced here from [50], it shows the relationship between the minimum detectable changes in frequency and the frequency at which they occur.

Table 6.1: The minimum detectable changes of frequency in Cents for Sine Waves

		Sensation Level (dBA)										
		5	10	15	20	30	40	50	60	70	80	90
Frequency (Hz)	31	220	150	120	97	76	70					
	62	120	120	94	85	80	74	61	60			
	125	100	73	57	52	46	43	48	47			
	250	61	37	27	22	19	18	17	17	17	17	
	550	28	19	14	12	10	9	7	6	7		
	1000	16	11	8	7	6	6	6	6	5	5	4
	2000	14	6	5	4	3	3	3	3	3	3	3
	4000	10	8	7	5	5	4	4	4	4		
	8000	11	9	8	7	6	5	4	4			
	11700	12	10	7	6	6	6	5				

A "cent" is defined as $\frac{1}{1200}$ of an octave [50], thus it represents a $\frac{2}{1200}$ change in frequency or 0.17%. The minimum detectable interval is 3 cents or a 0.5% total change in frequency.

6.3.3 Loudspeaker System Design

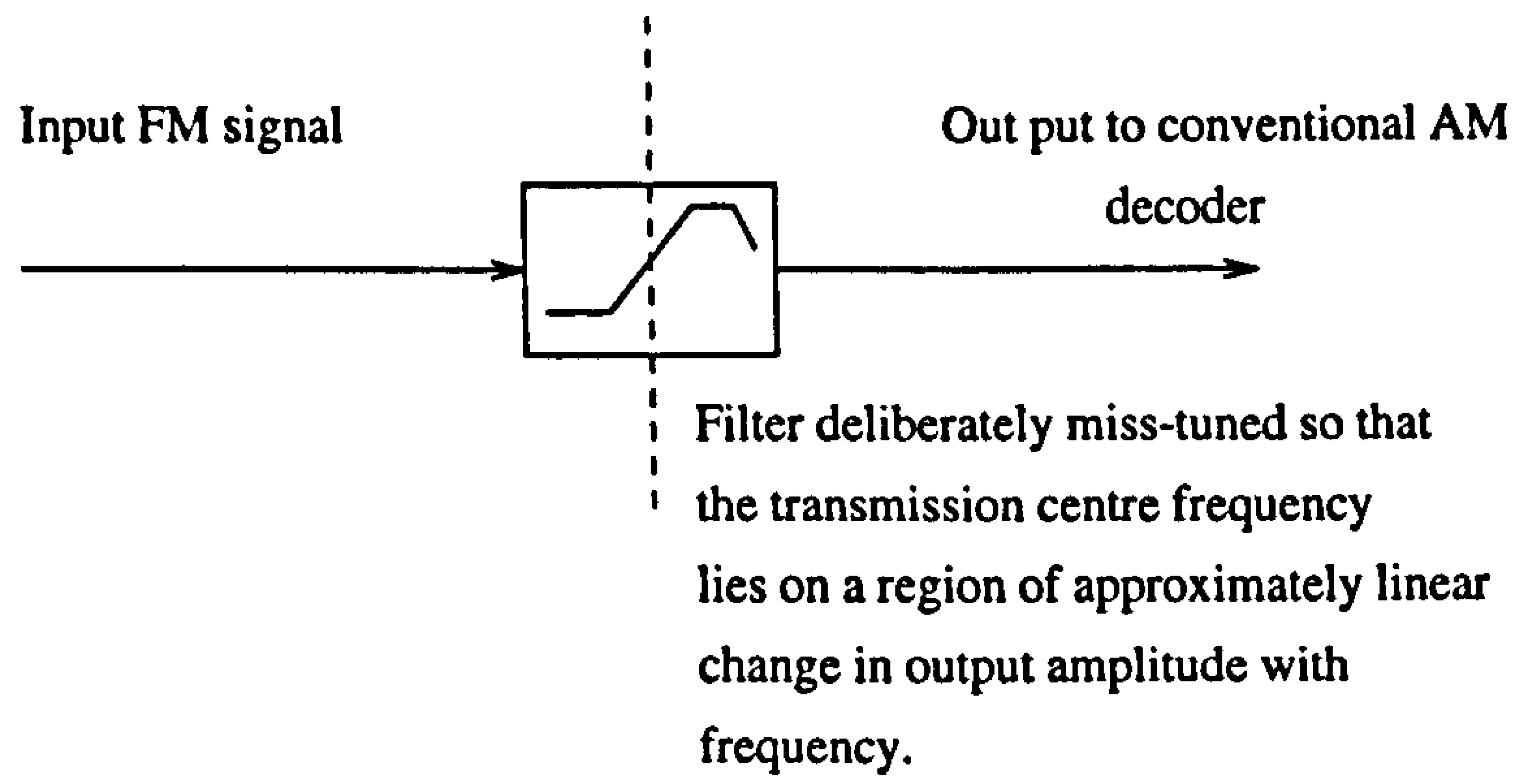
There are many trade-offs when designing a HiFi speaker, one of which is that it would be desirable both from economic and stereo image perspectives to use single drivers; since these best approximate a point source. In practice it has been found essential to use at least two drivers, sometimes many are used. Of the current production speakers by far the most common are those equipped with a relatively small diameter LF driver and a magnetic soft dome HF unit. In this combination the LF unit is expected to reproduce signals up to something like 3kHz. The large bandwidth and relatively small size of the LF unit will tend to exacerbate the Doppler distortion effect.

6.3.4 AM/FM Encoding and Decoding

Doppler distortion is an FM type effect and thus to detect and measure this effect is akin to FM decoding for radio work, provided that there are only two signal components in the test signal.

FM detection can be accomplished by a number of different methods, perhaps one of the simplest is to make use of a filter that is tuned slightly off the transmission frequency, so that as this changes it will produce a linear change in the amplitude of the output signal. The resulting signal is detected as for an AM transmission. This method will inherently also reproduce any AM component that might coexist with the desired FM component and thus the noise performance is not very good.

Figure 6.1: Double Side Band-Suppressed Carrier (DSB-SC) Demodulation

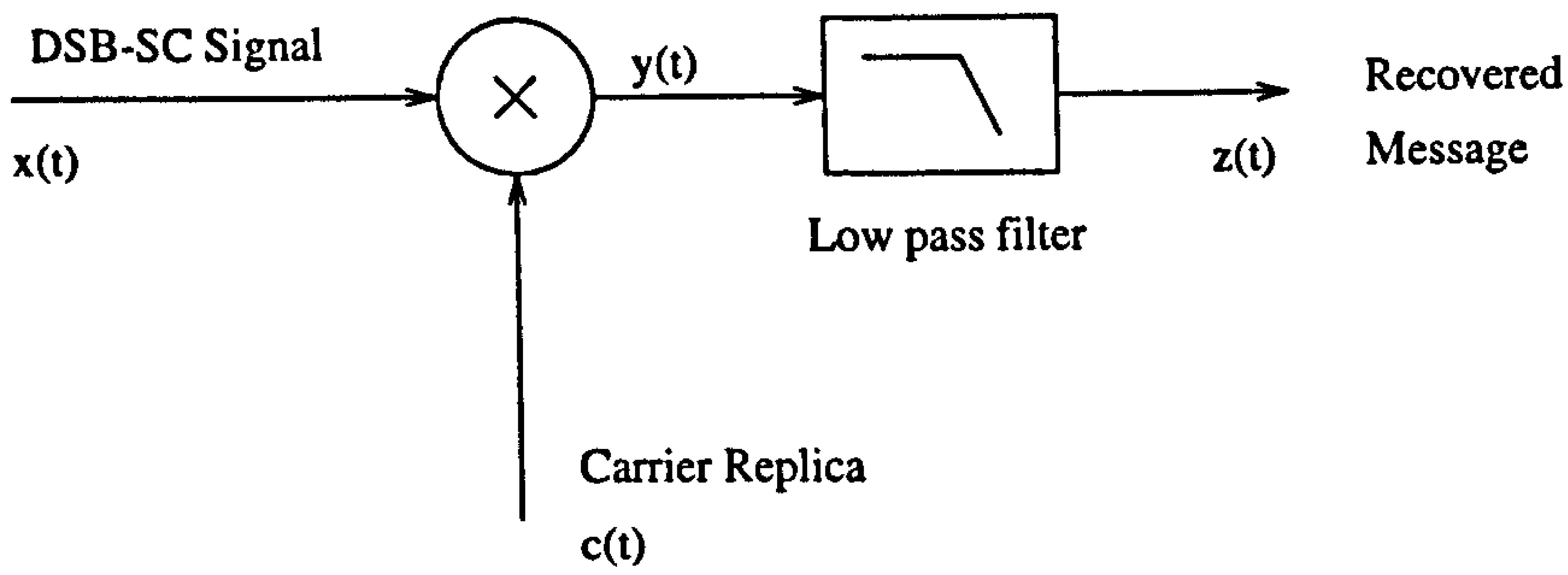


More sophisticated equipment tends to make use of a Phase Locked Loop (PLL), which can be configured to yield a linear output for a change in driving frequency. This class of device operates by performing a phase comparison between a carrier replica and the signal received.

There are many other methods of decoding FM but what was required was a methods that could easily and accurately be accomplished digitally.

The concept of a DSB-SC discriminator is used in the development and thus it is appropriate to review the associated theory here. If we take a conventional DSB-SC discriminator, [48].

Figure 6.2: DSB-SC Demodulation



Given that the signals are defined as:

$$x_{(t)} = A_m \cos \omega_m t \; A_c \cos \omega_c t \tag{6.2}$$

$$c_{(t)} = A_c \cos \omega_c t \tag{6.3}$$

Then the result of the multiplication will be:

$$y_{(t)} = x_{(t)} c_{(t)} \tag{6.4}$$

$$= A_m \cos(\omega_c t) \; A_c^2 (\cos \omega_c t)^2 \tag{6.5}$$

$$= A_m \cos(\omega_m t) \; \frac{A_c^2}{2} [1 + \cos 2\omega_c t] \tag{6.6}$$

$$= \frac{A_c^2}{2} A_m \cos(\omega_m t) + A_m \cos(\omega_m t) \; \frac{A_c^2}{2} \cos(2\omega_c t) \tag{6.7}$$

After the filter only the lower frequency component remains which is the recovered message. One of the problems that makes it difficult to apply this method for radio is that the carrier replica has to be of the correct frequency *and* phase. If there is a phase error in the carrier replica then the results are as follows:

$$\hat{c_{(t)}} = A_c \cos(\omega_c t + \Phi) \tag{6.8}$$

$$y_{(t)} = x_{(t)} \hat{c_{(t)}} \tag{6.9}$$

$$= A_m \cos(\omega_m t) A_c \cos(\omega_c t) \; A_c \cos(\omega_c t + \Phi) \tag{6.10}$$

$$= A_m \cos(\omega_m t) \; \frac{A_c^2}{2} [\cos \Phi + \cos(2\omega_c t + \Phi)] \tag{6.11}$$

$$= \frac{A_m A_c^2}{2} \cos(\omega_m t) \cos \Phi + \frac{A_m A_c^2}{2} \cos(\omega_m t) \cos(2\omega_c t + \Phi) \tag{6.12}$$

These results will be important in later sections in this chapter.

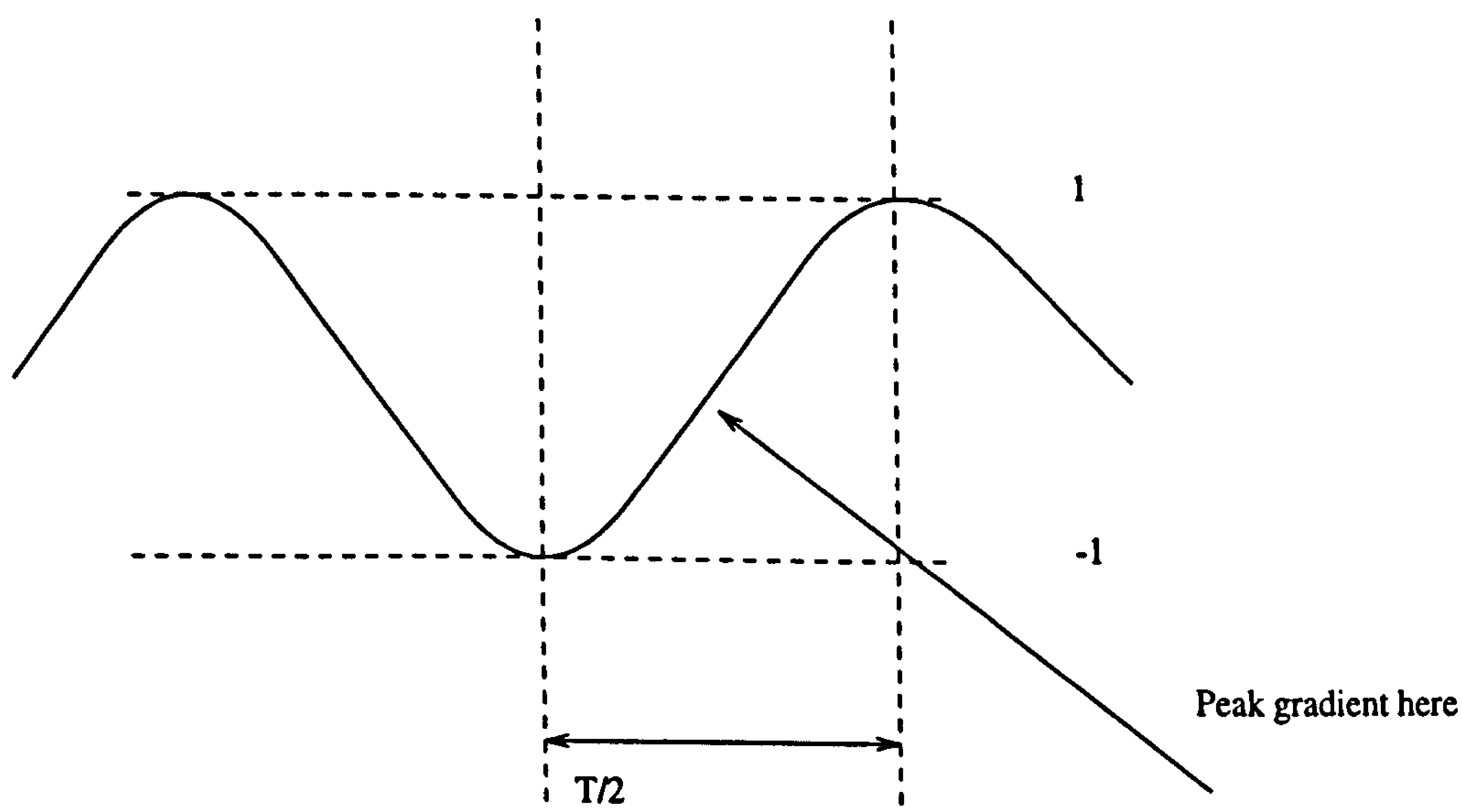
6.4 Design of Doppler Effect Compensator

6.4.1 Quantification of Doppler Distortion

The first stage of the distortion calculation is to make an estimate of the largest expected change in velocity of the loudspeaker cone, that is the peak to peak value, since the change will not take the form of deviation from the norm but rather oscillation between two peak values. (It is not possible to make direct measurements on the test speaker since this requires equipment that is not available, fortunately it is not necessary either.) A good estimate can be made by considering the peak to peak excursion that might be expected when the speaker is driven with a sine wave at low frequency,

Take as an example a cone experiencing a maximum displacement of .006m from peak to peak at 100 Hz. The cone travels 0.006m in 1/200sec which is an average speed of 1.2m/s. If the driving signal is a sine wave the peak speed of the cone is higher than this.

Figure 6.3: Peak Gradient of a Sine Wave



$$disp = \frac{3}{1000}\sin(100 * 2 * \pi * t) \tag{6.13}$$

$$vel = \frac{3}{1000}\cos(100 * 2 * \pi * t) * 100 * 2 * \pi \tag{6.14}$$

$$max(vel) = \frac{3 * 100 * 2 * \pi}{1000} \tag{6.15}$$

$$= 1.9 \tag{6.16}$$

In this case it is 1.9m/s. Given that the cone will be travelling in the opposite direction on the other half of the cycle the peak difference in velocity is double this. Using the formula:

$$\tilde{f} = \frac{c}{c - v} f \tag{6.17}$$

This translates to approximately ±1% error in reproduced frequency for all other signal components. This value of 2% total error represents four times greater than the minimum detectable interval for "normal" people. The assumed displacement is representative of a speaker operating at maximum power, except perhaps in the case of "long throw" types where the larger permissible excursion will cause the maximum deviation to be even greater.

To summarize: With a normal system operating at maximum power and reproducing a strong bass component it is probable that Doppler distortion will be audible. If the voltage drive to the speaker were reduced by a factor 4, which is an approximate power reduction of a factor of 8, the Doppler distortion would be expected to become inaudible.

In practice, the regular nature of the velocity signal may serve either to accentuate the effect since beat-like qualities or tend to mask the deviation since there will be LF signals present at the "beat" frequency. Practical observation would be required to clarify this point. Given the nature of the superFi market even systems that provide small improvements may be marketable. Doppler type distortion has proved to be inaudible with the poor grade equipment available to this project which means that a subjective appraisal has not been possible.

Similar effects to those described are obviously going to occur at the HF unit and recording microphone. Due to their nature these units inherently operate at much lower diaphragm velocities and thus the Doppler effects are not likely to be audible; the effects in the LF driver are not that far above the audible threshold.

The calculations are based on the listener being on the speaker's axis, so that the cone velocity is effectively the relative velocity of the listener and the cone. Simplistically speaking, if the listener were positioned parallel to the cone there would be no (almost) no relative velocity and thus the Doppler effect would not be expected to manifest. In practice the situation is probably more complicated. For instance it is reasonable, at low frequencies, to consider a hemisphere on an infinite plane as a basic model for a loudspeaker with the radius of the hemisphere changing to simulate the movement of the cone. In this case the radiation off axis would be in the direction of the velocity and thus Doppler effects should be expected. A better model is a flat disk in an infinite plane, which produces in a very similar result except that the intensity of the signal off axis depends on the angle and the frequency of the signal. Thus it would seem that the Doppler effect should be expected to be as strong off-axis as on, at least as a first approximation. More accurate information would require a detailed analysis of the near field to a loudspeaker. Even if the Doppler effect were weaker or non-existent off axis this has little bearing on the practical application of this concept since listeners are positioned on axis or nearly so to avoid the (much stronger) effects of the non-ideal polar output of real loudspeakers.

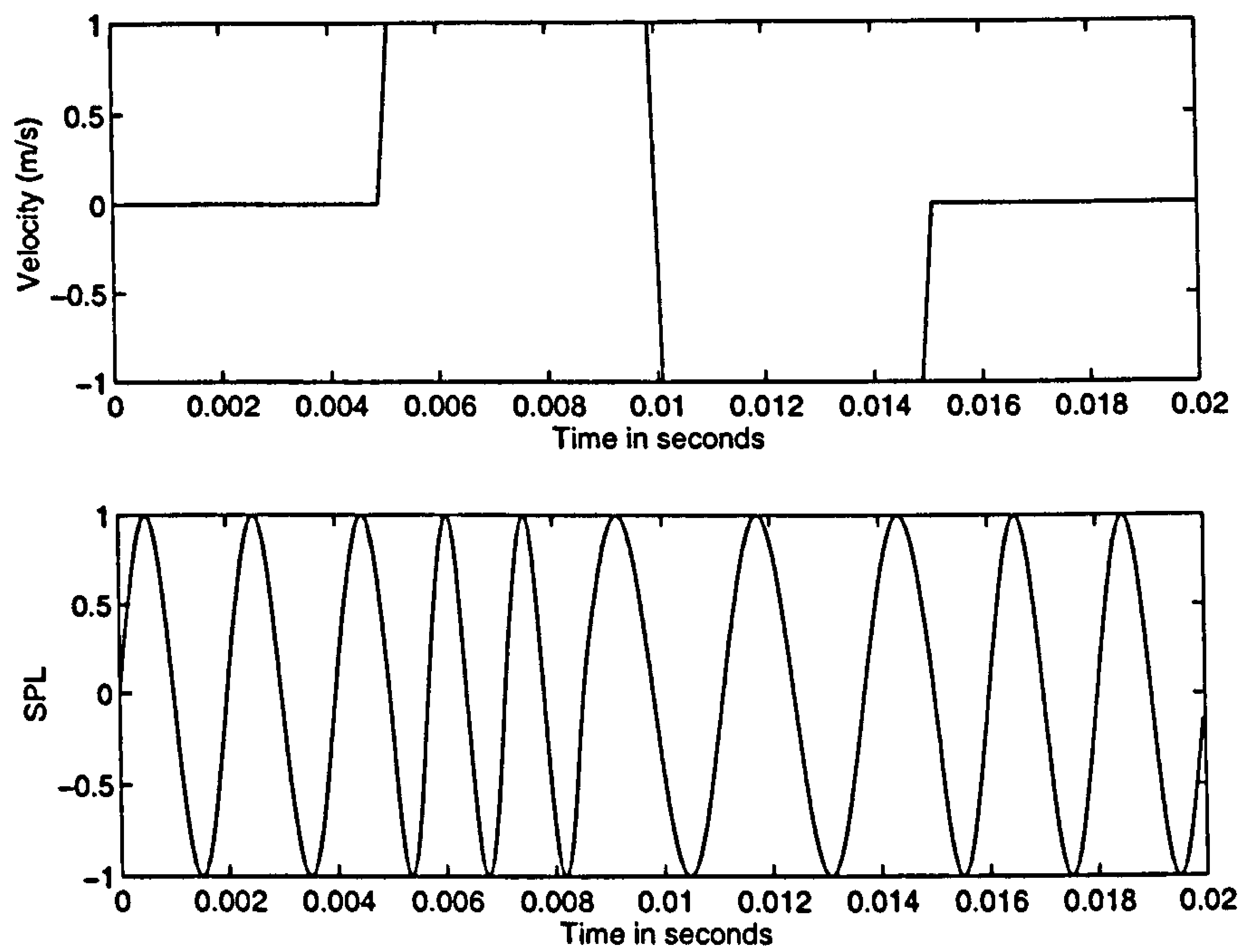
The Doppler effect would have an interesting effect on the polar response of the speaker. The polar response is caused, simplistically speaking, by the summation of all possible paths from the radiating surface to the point of interest. The Doppler effect will cause the amplitude of the result of a single signal path to depend not only on the path length (a delay) but on the velocity (as a result of other signals only) of the cone at the time of emission. Thus the polar response when emitting two or more frequencies would be time varying.

6.4.2 Implementation

The calculations in the previous section revealed that there will only be a need to compensate for Doppler type distortion for the LF driver. The HF driver will be operating at quite a different and much lower velocity and thus this signal should not pass though the compensation.

It is essential consider how exactly the Doppler effect distorts the signal. If the graph below is considered it can be seen that the effect is to "squash" and "stretch" the signal depending on the velocity at the time of emission. While this has the effect of changing the frequency as presented earlier it also causes apparent changes in the length of time the source spends at any velocity from the point of view of the listener. (Note that the velocities used to create the graph are very large so that the effects are clearly visible.

Figure 6.4: A Doppler Distorted Signal



It might seem a little odd that the perceived signal frequency can increase when the emitted frequency is constant, it seems to indicate that "information" is reaching the listener faster than it is being emitted. In fact that is exactly the case, it is possible only because of the "stored" information already in transit. The frequency cannot be indefinitely altered for finite distances. In the previous figure the effect can be seen clearly, while the velocity changes for roughly even intervals to a number of different values the resulting signal is perceived to be at a higher frequency for less time than it is at a lower value.

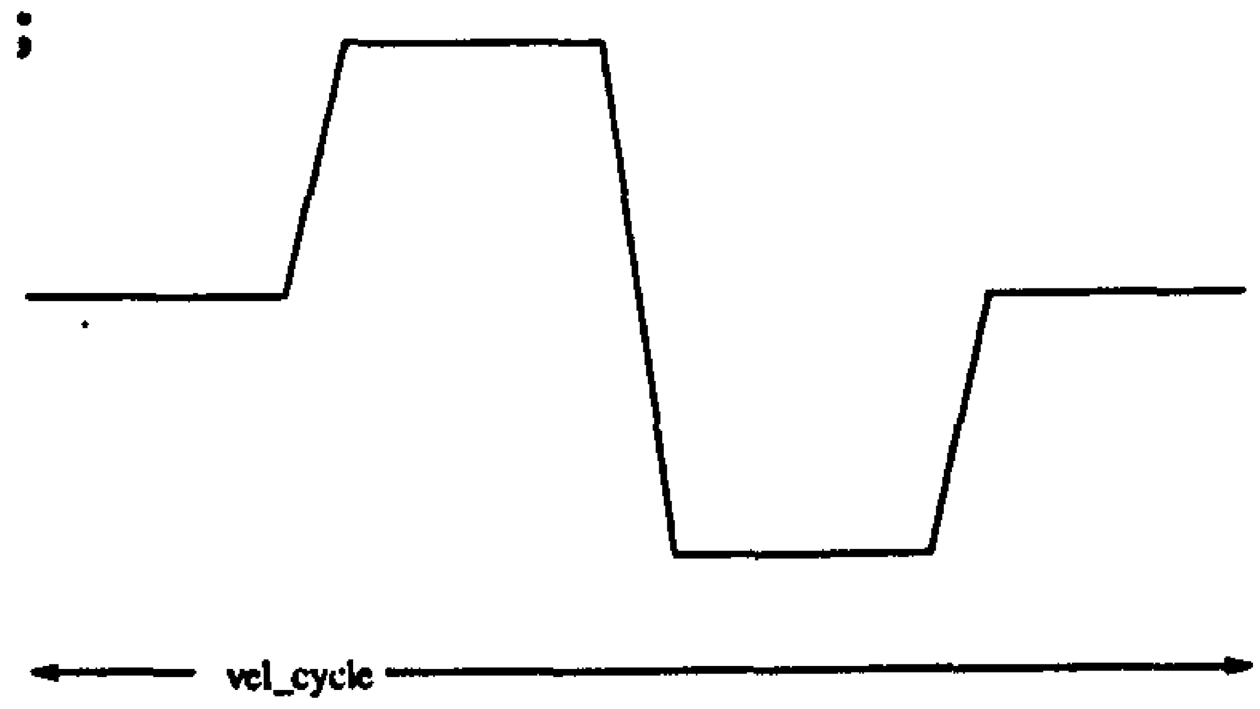
In order to correct for Doppler distortion it will be necessary to make a process that has the opposite effect on the signal. The normal methods of filter design etc. are of no use here; no linear filter will correct for this type of distortion. The distortion is, however, predictable and no information is lost (unlike clipping for instance) and thus the damage appears reversible.

Simulating Doppler Distortion

The first stage towards producing a corrective filter will be to produce a good simulation of the effects it is intended to counter. For the purposes of simulation it is sufficient at first to use an arbitrary, pre-defined, source (cone) velocity. A simple velocity signal was constructed using very large peak values (100m/s) so that the effects could be seen easily.

This code creates a sequence of velocity changes and then copies this repeatedly into *vel* to make the velocity signal. In between the changes in value finite slopes have been added, so that the result is a little more realistic.

```
vel_cycle = [zeros(1,intvl-trans/2)       linspace(0,1,trans) ...
             ones(1,intvl-trans)          linspace(1,-1,trans) ...
             (-1).*ones(1,intvl-trans)    linspace(-1,0,trans) ...
             zeros(1,intvl-trans/2)].*Vcont;
for i=1:Tmax*50,
   vel = [vel vel_cycle];
end
```



The next stage is set up a vector of zeros as the new time axis t_axis, this is done largely to ensure reasonable performance from Matlab.(Failing to allocate space in advance forces Matlab to perform a memory allocation and copy every time an additional variable is added to the end of a vector, with the attendant performance costs.) The following line is necessary or Matlab will not allow the interpolation process in the loop. In the loop the new time axis calculated on the basis of the previous time and velocity values. (The interpolation is necessary since *vel* is not continuously defined.) It is noted that the process may suffer from error growth if long tests were to be carried out. Since this did not occur no protection has been included.

```
t_axis = zeros(1,length(t)-buf/2);
t_axis(2) = 1/44000;
for i=3:(length(t)-buf/2),
   t_axis(i) = t_axis(i-1)+c/((c-interp1(t,vel,t_axis(i-1)))*44000);
end
```

Looking at the loop in more detail it can be seen that the next value of t_{axis} is the previous value with an adjustment added based on the velocity at time $t_{axis}(i-1)$. The equation below is the same function expressed in a more familiar mathematical notation. Note that *vel* is considered as a continuous function instead of the interpolation.

$$t\_axis_i = t\_axis_{i-1} + \frac{1}{44000}\frac{c}{(c - vel_{(t\_axis_{i-1})})} \tag{6.18}$$

The factor $c/(c - vel)$ originates from basic Doppler theory, as presented in the subject revision.

Normal and Doppler distorted copies of the signal are made. Note that both of these signals should be plotted against t despite the distorted one being created using t_axis. For large velocities it is easy to see that the process behaves as expected, the graph presented earlier of a "Doppler Distorted Signal" was created using this method.

```
sig = sin((Fsig*2*pi).*t);
Ssig = sin((Fsig*2*pi).*t_axis);
```

Simple Doppler Distortion Correction

It is not difficult to see that it should be possible to reverse the distortion by applying a modified version of the procedure used during it's simulation. There is a slight additional complication, the signals used in the simulation were continuously defined but here the input signal is sampled. The only option is to re-sampled against the corrected time axis, which will involve interpolation of the signal.

The interpolation process will inherently introduce distortion but this is expected to be minimal since the process will only be applied to the signal destined for the LF driver; this implies that there will be many samples defining even the higher frequency waveforms. If compensation were required for the HF driver the problem of distortion would potentially be far more serious. At this stage only linear interpolation is used. If there were to be a subjective problem with distortion caused by the interpolation then a more complicated scheme could be used at the cost of an additional processing power requirement.

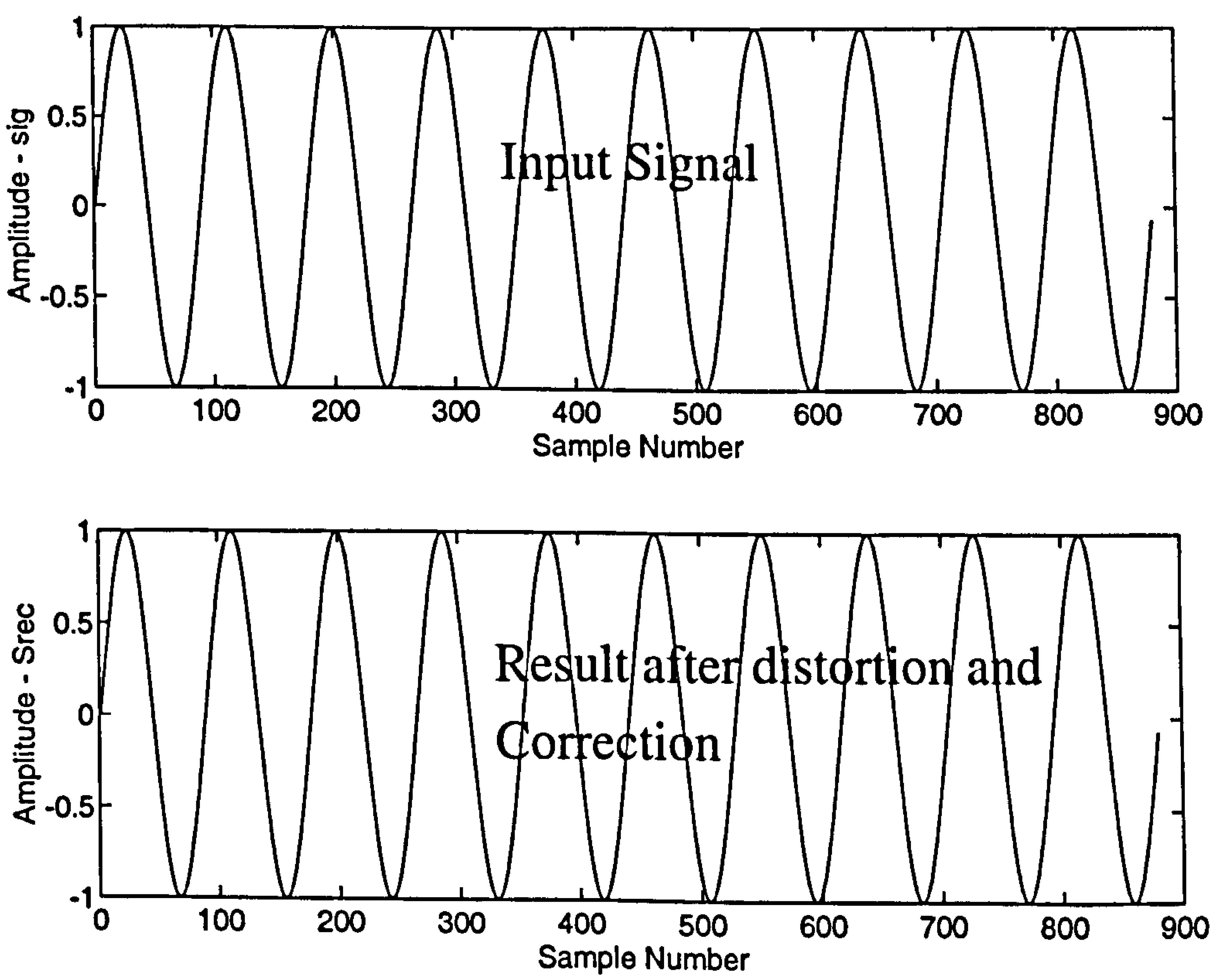
```
t_axis_r = zeros(1,length(t)-buf);
t_axis_r(2) = 1/44000;
for i=3:(length(t)-buf),
   t_axis_r(i) = t_axis_r(i-1)+((c-vel(i-1))/(c*44000));
end
```

```
Srec = interp1(t(1:length(Ssig)),Ssig,t_axis_r);
```

Where *Ssig* is the Doppler distorted signal from the earlier simulation and *Srec* is the corrected signal. Note that in this case there in no need to perform an interpolation when calculating t_axis_r since the velocity required is with respect to the new time axis, which in this case is t.

The result of this process is an accurate approximations to the original signal, the results can be seen in figure 6.5. The original signal is given in the upper plot and the distorted and corrected version in the lower. The corrected version has apparently been accurately restored and is now very similar to the original signal. The level of distortion applied can be seen in figure 6.4. In the real system the compensation will have to occur before the distortion effect instead of after, as has been the case in this demonstration. Reversing the order of the correction and distortion processes has no effect on the final result.

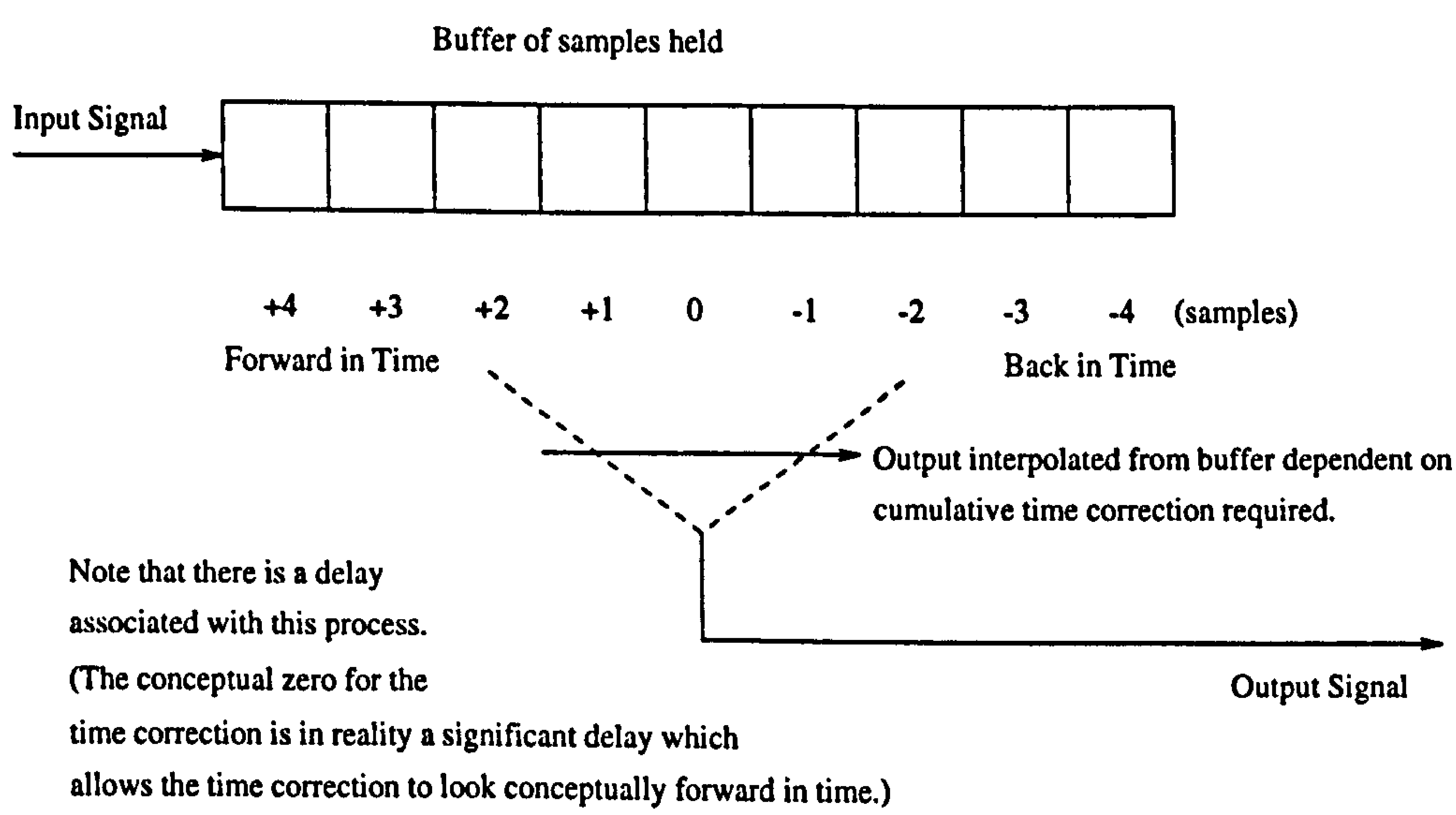
Figure 6.5: Signal and Result after Distortion and Correction



Correction Relative to Current Time

The simplistic solution presented above could not be used in any real system since it uses the entire signal and time vectors, which would rapidly become unworkable large for any realistic signal length. In practice there is no need for all of this information to be held and thus the next stage is to produce a system that will operate relative to the current point in time. The larger signal lengths will also require protection against error growth in the cumulative operations. The conceptual operation is shown in the figure below:

Figure 6.6: Conceptual Method for Doppler Distortion Correction



The changes required are slight. If the current sample is arranged to be at the centre of a sufficiently long buffer and the time shift is calculated rather than the absolute time then the full signal and time vectors are no longer required. It is also necessary to add a little "negative" feedback to the calculation of the time-shift. The Matlab code to implement this is included in the appendices, section 9.8, a summary of it's operation is given below.

```
for each sample
   {
   update sample buffer
   *** calculate new time shift ***
   timeshift = previous time shift + additional correction for this sample
```

```
                    - small part of previous time shift.
  output = interpolated from sample buffer at time shift
  }
```

The altered code performs very much like the previous version provided that the centralizing value is small enough. (There is no reason why this value should be large since it is only to compensate for error growth in calculation, which will probably be very small.) There is one remaining limitation, the compensation still relies on a separate velocity vector.

System Design

At the conclusion of the previous section a complete "resampler" had been made but this still required the input of a velocity signal separate from the main data stream. There are two different approaches that can be taken to this problem; either the velocity can be measured by some method, for instance by making use of the back EMF as discussed in the previous chapter, or an estimate can be made on the basis of a model of the speaker. Since using a velocity estimate requires no additional hardware this method was chosen, although there is no doubt that the performance would improve if the velocity could be measured accurately. If the system were already equipped with the back EMF control system considered in the previous chapter then it makes sense to use this data, if only to tune the model. There is no delay inherently associated with the *output* of the re-sample process and thus it is the current value of the cone velocity that is required by the re-sampler.

With a two driver speaker system the signal to the HF driver will not require compensating; the Doppler type distortions from this unit would be at very much lower levels due to the lower cone velocity.

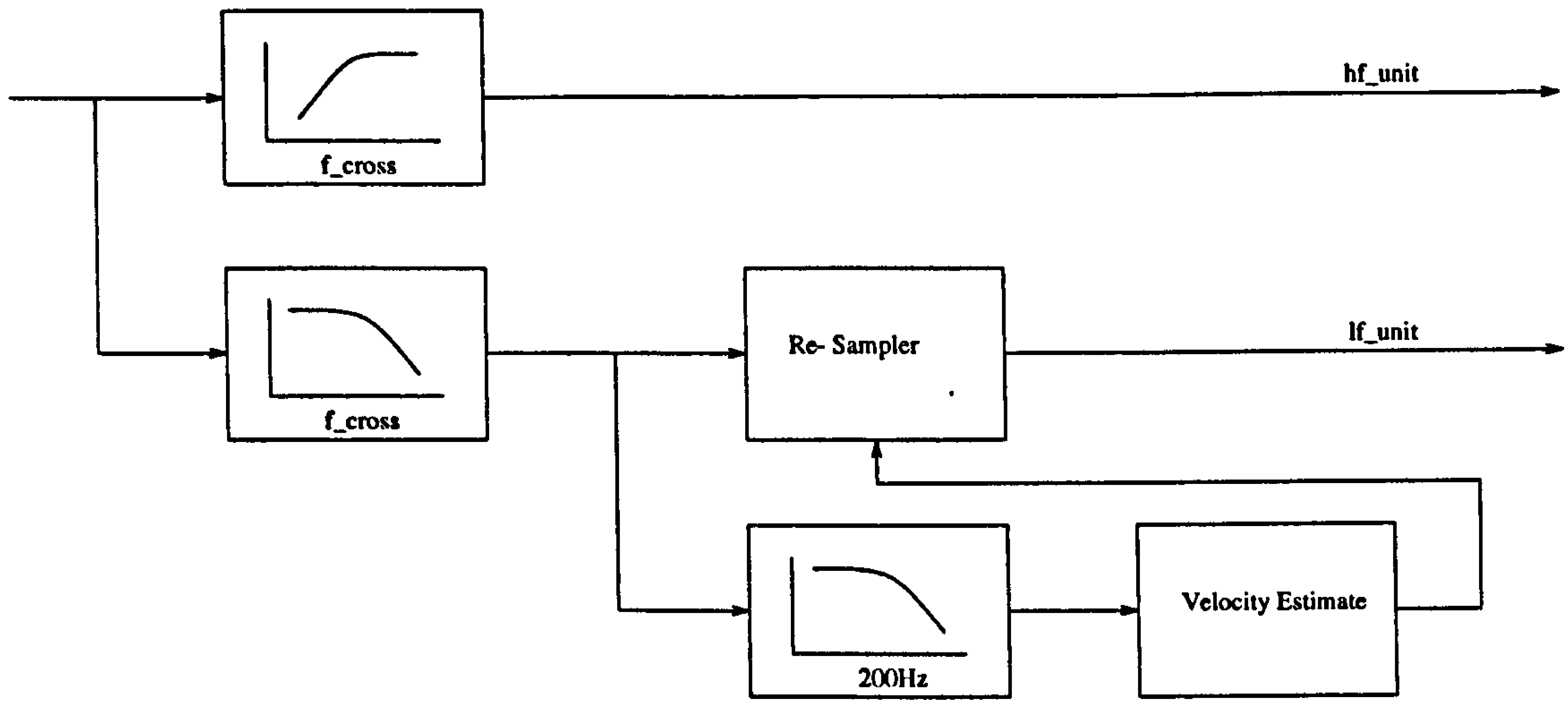
It is well known that for low frequencies the bulk parameter models for speakers can provide good results but the accuracy of this method reduces rapidly above the first break-up mode of the cone. If a model based method for determining the velocity were to be chosen then it would be important to limit the range of frequencies into the model to those for which it was valid. It is not as important to estimate the velocity at higher frequencies since the peak values expected fall at above the system resonance, which normally occurs at between 50 and 100Hz. The frequency at which the first breakup mode occurs is very much dependent of the design of the driver in question, but is normally of the order of 700Hz [12].

It is essential that the corrected signal is only be sent to the base driver since any portion of the signal reproduced by the HF unit will be distorted by inappropriate correction. There are two ways to ensure this, one is to split the signal digitally, carry out the required processing and then use two DACs per channel and bi-wire the speakers. There are numerous advantages to this method, not least that the crossover is now performed in the digital domain which allows near ideal filters both in terms of amplitude and phase. (It is possible to make digital filters with constant group delay, ie. effectively zero phase change, that are of very high order and are guaranteed matching. Digital filters do not drift and can be reproduced exactly. Analogue filters are inherently imperfect since real components are inevitably exhibit parasitic effects, are often associated with rapid phase changes, must be expected to drift with changing operating conditions/time and can normally only be matched to the limit of the component tolerances/accuracy of the initial tuning.) Alternatively the signal could be re-combined after the compensation to be split again in the speaker's cross-over network. The inherent deficiencies of the passive networks fitted to loudspeakers will mean that the compensated signal cannot be confined to the LF driver as effectively.

Regardless of the method chosen any filters that exist between the compensation unit and the speaker are very important if the velocity estimate is to be made on the basis of a model. Provided that the filters are implemented in the digital domain there is the opportunity to use the FIR structure which are ideal for this task since they can be designed to have a linear phase characteristic. A linear phase shift is simply a delay and is therefore easy to compensate for when making the velocity estimate. (It is possible that the improved characteristics of a cross-over implemented in the digital domain may produce a significant improvement in perceived sound quality.) For the purposes of this demonstration digital filters have been used.

The figure below describes the conceptual operation of the system. (Some implementation details have been omitted for clarity, full details are given in figure 6.10.)

Figure 6.7: Conceptual System for Doppler Distortion Correction



It is worth considering the consequences of error in the velocity estimate. For instance if the velocity estimate were to be say 30% too small then the correction would fall short of the ideal but would still provide a useful improvement. A phase error is slightly more serious, this will cause the correction to be stretching the signal at times when it should be squashed etc. Even in this case though the damage is not very serious provided that the error is quite small since just after the zero crossing the velocity values are small which in turn means that the incremental time shift values are small leading to only a relatively small error over all in time shift.

Extending these concepts into a more rigorous mathematical framework, the apparent altered sample time due to the Doppler effect is given by:

$$\hat{T}_s = T_s \frac{c}{c - vel} \tag{6.19}$$

During the compensation the reverse process is carried out so the the sample time interval is effectively consistent.

$$\bar{T}_s = \hat{T}_s \frac{c - vel}{c} \tag{6.20}$$

Carrying out the obvious substitution:

$$\bar{T}_s = T_s \frac{c}{c - vel} \frac{c - vel}{c} \tag{6.21}$$

$$= T_s \tag{6.22}$$

Of course, in practice there will be an element of mis-match between the calculated and the actual velocities, $\hat{vel}$ and vel respectively. In that case the we have to distinguish between the actual velocity and the approximation.

$$\bar{T_s} = T_s \frac{c}{c - vel} \frac{c - \hat{vel}}{c} \tag{6.23}$$

Thus the effect of a mismatch between the measured and the actual velocity will be a frequency modulation distortion, but this time according to a function given by the mismatch in the two functions *vel* and $\hat{vel}$. If we take *vel* and $\hat{vel}$ to be:

$$\begin{aligned} vel &= \sin(\omega t) && (6.24) \\ \hat{vel} &= A \sin(\omega t + \Phi) && (6.25) \end{aligned}$$

Where A and Φ represent the amplitude and phase error respectively. Considering their net effect on T_s

$$\frac{\bar{T_s}}{T_s} = \frac{c - A\sin(\omega t + \Phi)}{c - \sin(\omega t)} \tag{6.26}$$

If it is assumed that the filter will be beneficial while the maximum deviation of the ratio $\frac{\bar{T_s}}{T_s}$ from unity is less than that of $\frac{\hat{T_s}}{T_s}$ then the following is true:

$$\left|1 - \frac{c}{c - \sin(\omega t)}\right|_{MAX} > \left|1 - \frac{c - A\sin(\omega t + \Phi)}{c - \sin(\omega t)}\right|_{MAX} \tag{6.27}$$

Looking at the left hand side of the expression above:

$$\begin{aligned} \left|1 - \frac{c}{c - \sin(\omega t)}\right|_{MAX} &= \left|\frac{c - \sin(\omega t) - c}{\sin(\omega t) - c}\right|_{MAX} && (6.28) \\ &= \left|\frac{\sin(\omega t)}{\sin(\omega t) - c}\right|_{MAX} && (6.29) \\ & \qquad\qquad given \quad c >> \sin(\omega t) && (6.30) \\ &\approx \left|\frac{\sin(\omega t)}{c}\right|_{MAX} && (6.31) \\ &\approx \frac{1}{c} && (6.32) \end{aligned}$$

Similarly with the right hand side:

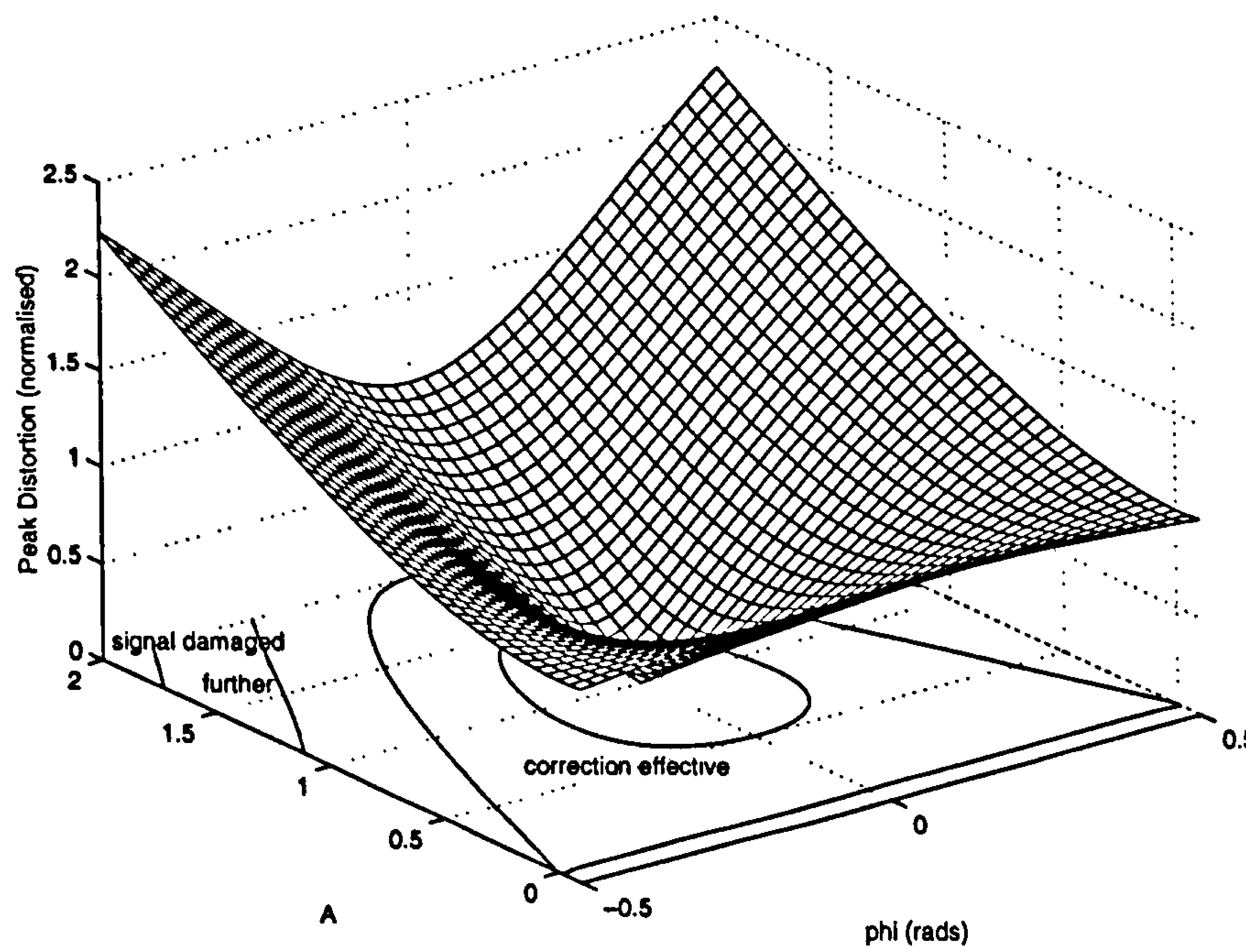
$$\begin{aligned} \left|1 - \frac{c - A\sin(\omega t + \Phi)}{c - \sin(\omega t)}\right|_{MAX} &= \left|\frac{A\sin(\omega t + \Phi) - \sin(\omega t)}{c - \sin(\omega t)}\right|_{MAX} && (6.33) \\ & \qquad\qquad given \quad c >> \sin(\omega t) && (6.34) \\ &\approx \left|\frac{A\sin(\omega t + \Phi) - \sin(\omega t)}{c}\right|_{MAX} && (6.35) \end{aligned}$$

Thus the inequality becomes:

$$\begin{aligned} \frac{1}{c} &> \left|\frac{A\sin(\omega t + \Phi) - \sin(\omega t)}{c}\right|_{MAX} && (6.36) \\ 1 &> |A\sin(\omega t + \Phi) - \sin(\omega t)|_{MAX} && (6.37) \end{aligned}$$

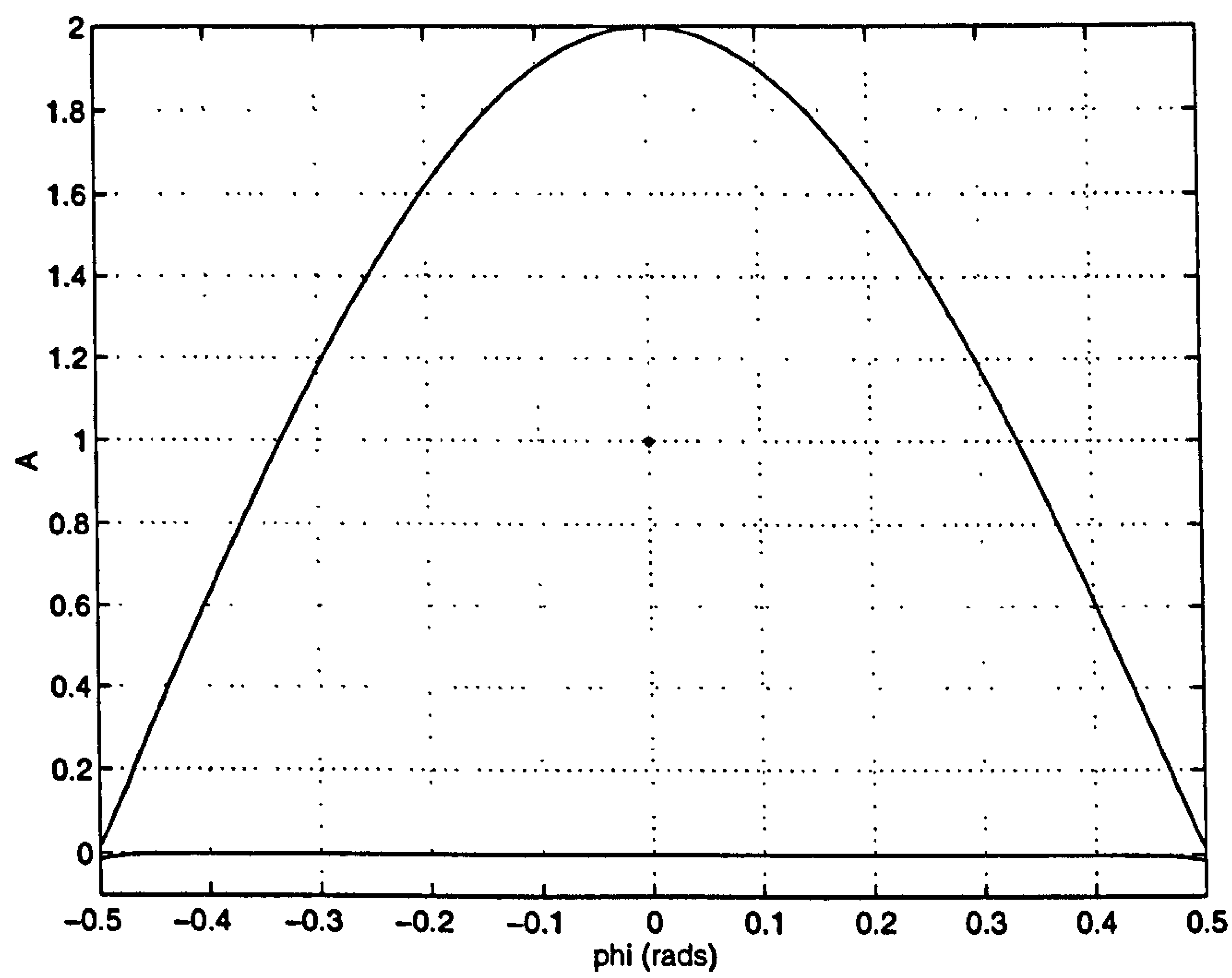
It is difficult to proceed further analytically and there is little reward since Matlab is able to provide a numerical solution quickly. A surface representing the right hand side is given in figure 6.8, where its value is below unity the filter is effective as defined by the initial assumption.

Figure 6.8: Normalized Peak Distortion



More important perhaps than the shape of the surface are the limiting values of A and Φ. Figure 6.9 is a single contour of the previous graph located at unity; if A and Phi placed the system at or beyond this line then the attempted correction would increase the distortion levels further. The dot in the centre indicates the optimum values.

Figure 6.9: Limit in Values of A and Φ for Correction



Note that some of the resulting distortion functions of time (the modulating functions) are more complicated and may possibly be audible at lower levels than the original Doppler distortion.

Extending this analysis further to include a velocity signal made up of a number of discrete sine waves. If we re-define vel as:

$$vel = \sum_{n=0}^{n=m} a_n \sin(\omega_n + \phi_n) \tag{6.38}$$

Thus $\hat{vel}$ is now equal to:

$$\hat{vel} = \sum_{n=0}^{n=m} A_n a_n \sin(\omega_n + \phi_n + \Phi_n) \tag{6.39}$$

That is, the velocity simulation is in error by a different A_n and Φ_n for each of ω_n. The question is what effect does this complication have on the distortion expected and the tolerance to error. To start with we take a simple case where the velocity is made up of two components:

$$vel = vel_1 + vel_2 \tag{6.40}$$

Using the previous analysis the new time interval will be:

$$\hat{T}_s = T_s \frac{c}{c - (vel_1 + vel_2)} \tag{6.41}$$

If instead we were to consider the new time interval by looking at the effects of vel_1 and vel_2 separately then we would be slightly in error:

$$\hat{T}_s \approx T_s \frac{c}{c - vel_1} \frac{c}{c - vel_2} \tag{6.42}$$

$$\approx T_s \frac{c^2}{c^2 - c\,vel_1 - c\,vel_2 + vel_1 vel_2} \tag{6.43}$$

$$\approx T_s \frac{c}{c - (vel_1 + vel_2) + \frac{vel_1 vel_2}{c}} \tag{6.44}$$

Given that vel_1, vel_2 are very much smaller than c the last term may be ignored since it will be small. Since this approximation can be accepted then it can be seen that it would be sufficient to consider the effects of vel_1 and vel_2 separately.

If we return to the more general definition of vel and $\hat{vel}$ and make use of the previous analysis of the acceptable distortion limits we have:

$$\prod_{n=0}^{n=m} |A_n a_n \sin(\omega t + \phi_n + \Phi_n) - a_n \sin(\omega t + \phi_n)|_{MAX} < \prod_{n=0}^{n=m} a_n \tag{6.45}$$

Which simplifies to:

$$\prod_{n=0}^{n=m} |A_n \sin(\omega t + \Phi_n) - \sin(\omega t)|_{MAX} < 1 \tag{6.46}$$

To summarize, provided that for each of vel the relevant $\hat{vel}$ is within the limits displayed in the graph developed during the previous analysis then the levels of distortion due to this effect will definitely be reduced. It is possible that even if the error exceeds this limit for some of vel the overall distortion levels may still be reduced. The graph in question may therefore be used as a reasonable indication of the acceptable error in determining the cone velocity.

Returning to the correction system design. It is necessary to consider the problem of self distortion, as explained earlier the Doppler effect does not occur with a signal consisting of a single sine wave. The compensation strategy, however, will have a velocity figure and thus will inevitably attempt to compensate for an effect that will not occur. In practice the velocity will typically be made up from many different components and thus the distortion caused to each component by this simplification will still be out-weighed by the potential gains.

There is a strategy that can be used to circumvent this problem: If the signal that is to be sent to the LF driver were split about some fixed frequency, say of the order of 400Hz then the LF components could be used to make a velocity estimate on which basis the higher frequency components could be corrected. (Below about 400Hz the sensitivity of the human ear to frequency shifts falls rapidly, as can be seen with reference to table 6.1 and above the resonant frequency the velocity falls rapidly and thus the contribution to the velocity made by signals above this frequency will be small.) 400Hz is also well below the expected first break-up mode and thus a model should provide a reasonably accurate estimate of the cone velocity.

A negative feedback loop has been added to the basic system to ensure that the resampler does not slowly drift off the end of its buffer. The addition is very simple, a small constant multiplied by the current position of the resampler in the buffer is subtracted from its location each sample. This should offer ample protection against computational error growth.

There is a delay associated with the re-sampling buffer that needs to be compensated for, as do the delays due to the FIR filters. The complete system is described in the two diagrams on the following pages, the latter being a detail of the "re-sampler".

Figure 6.10: System for Doppler Distortion Correction

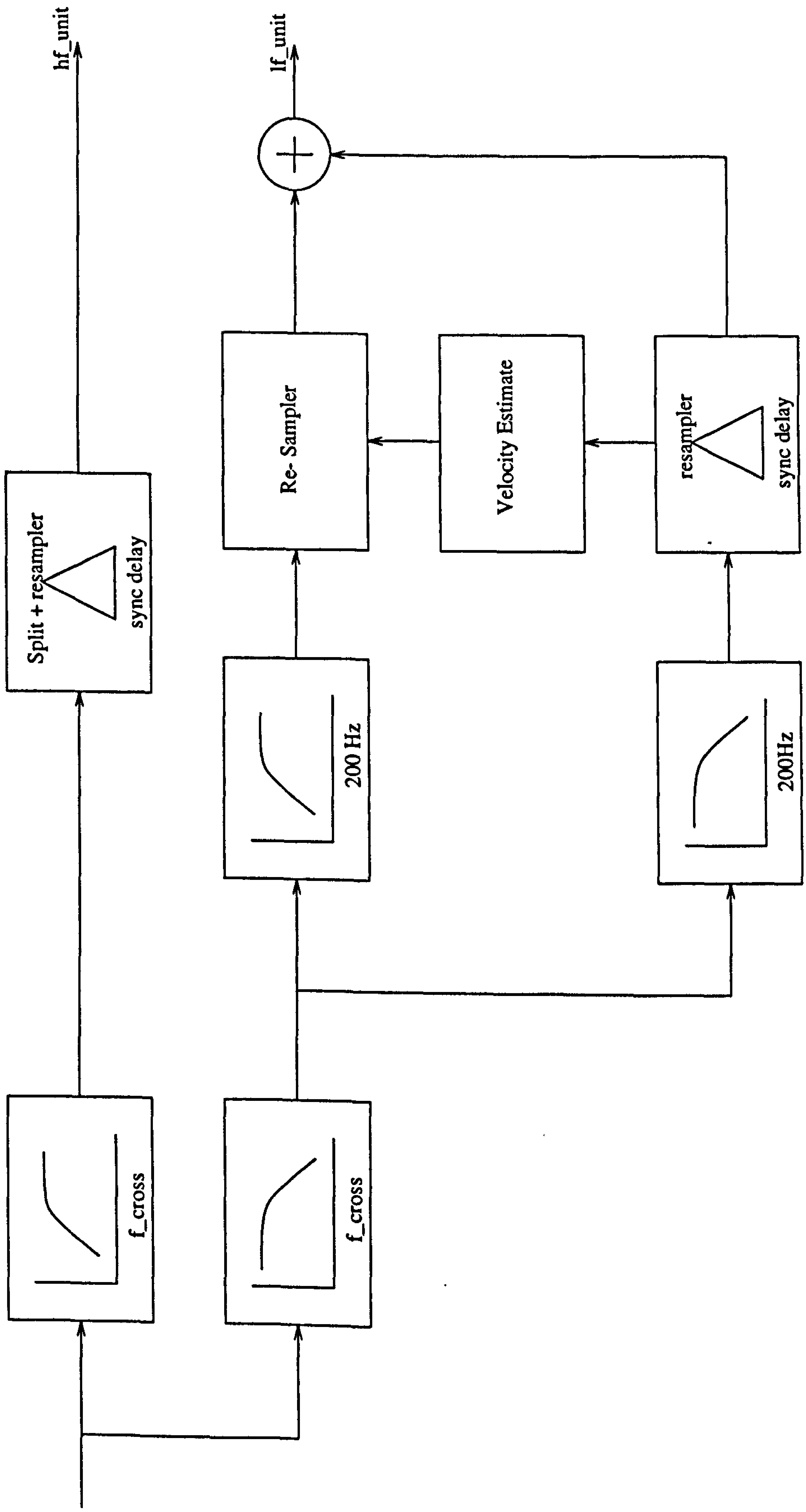
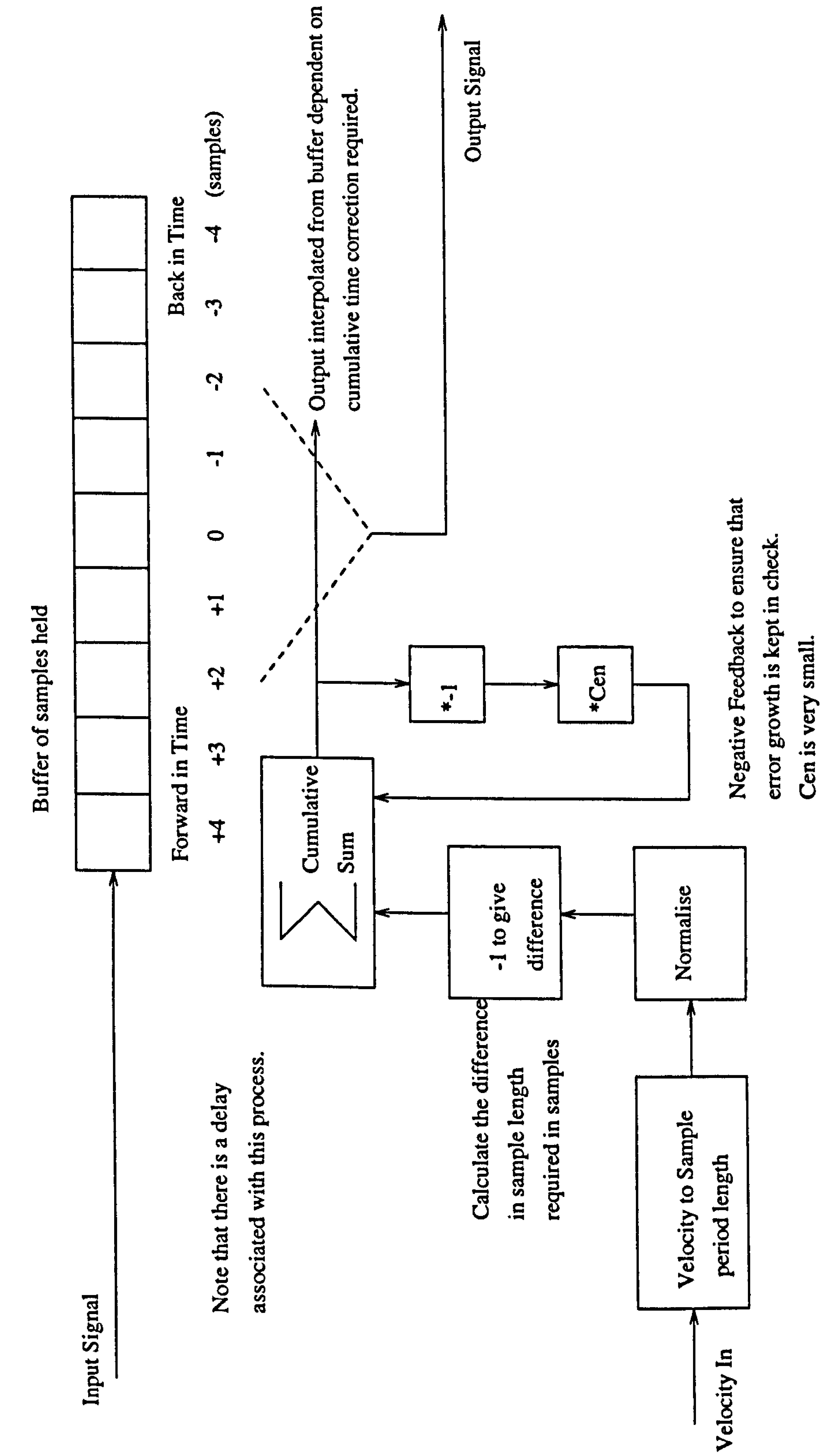


Figure 6.11: Detail of the “Re-sampler”



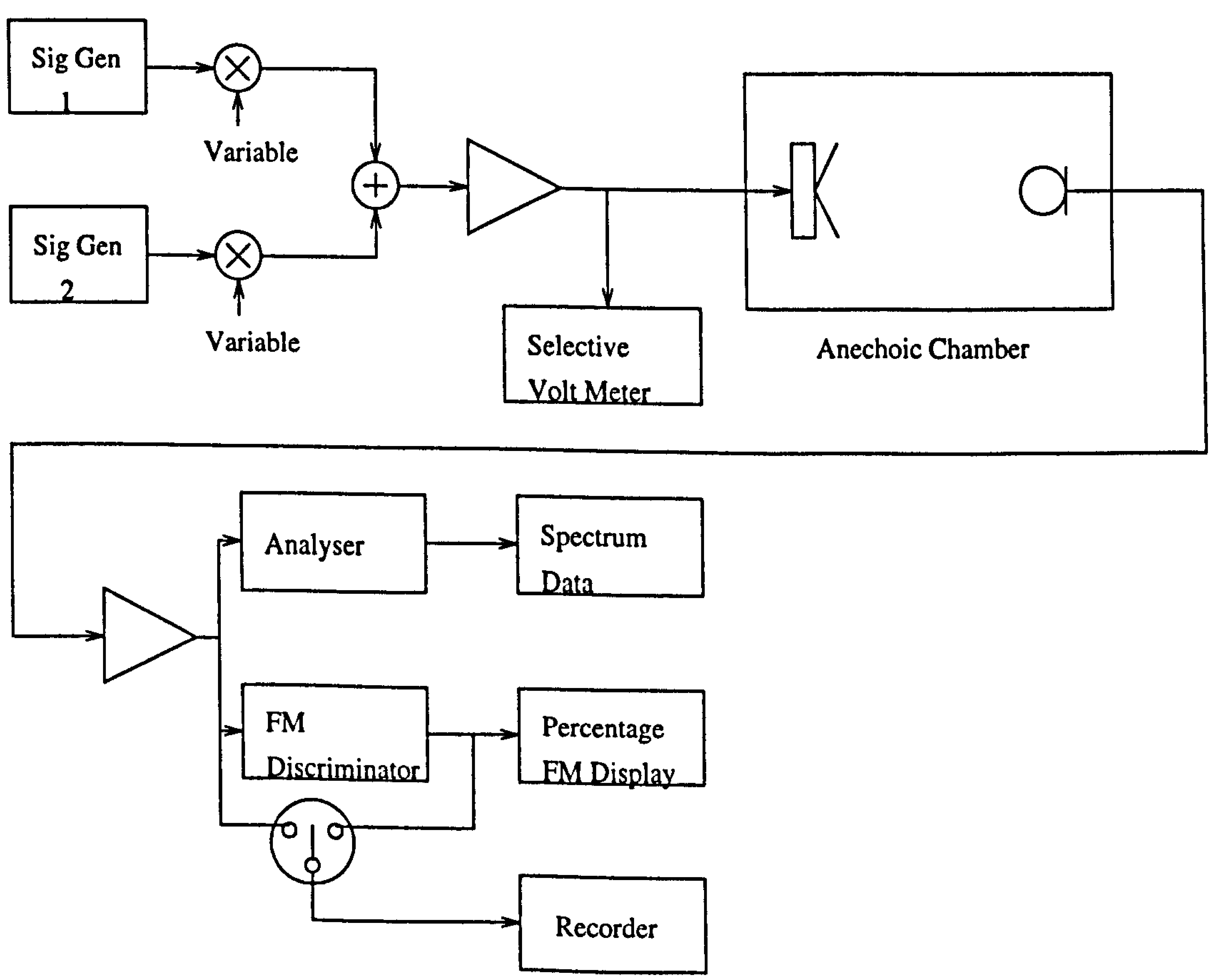
6.4.3 Testing

Initial tests revolved around correcting an audio signal and then attempting to discern the difference in the quality of the reproduced sound. These tests were not carried out on a formal basis since it appeared that it was not possible to detect any difference given the low quality of the reproduction equipment available. It should be understood that this concept is really intended for application to audiophile grade equipment and so this result came as no great surprise. Due to the limited finance available to this project buying suitable test equipment was not a realistic possibility therefore alternative methods were considered.

Since it has already been demonstrated that the concept works under simulation it was necessary to design a test that used a real loudspeaker. This complicated matters since the added noise and other distortions made the measurement difficult.

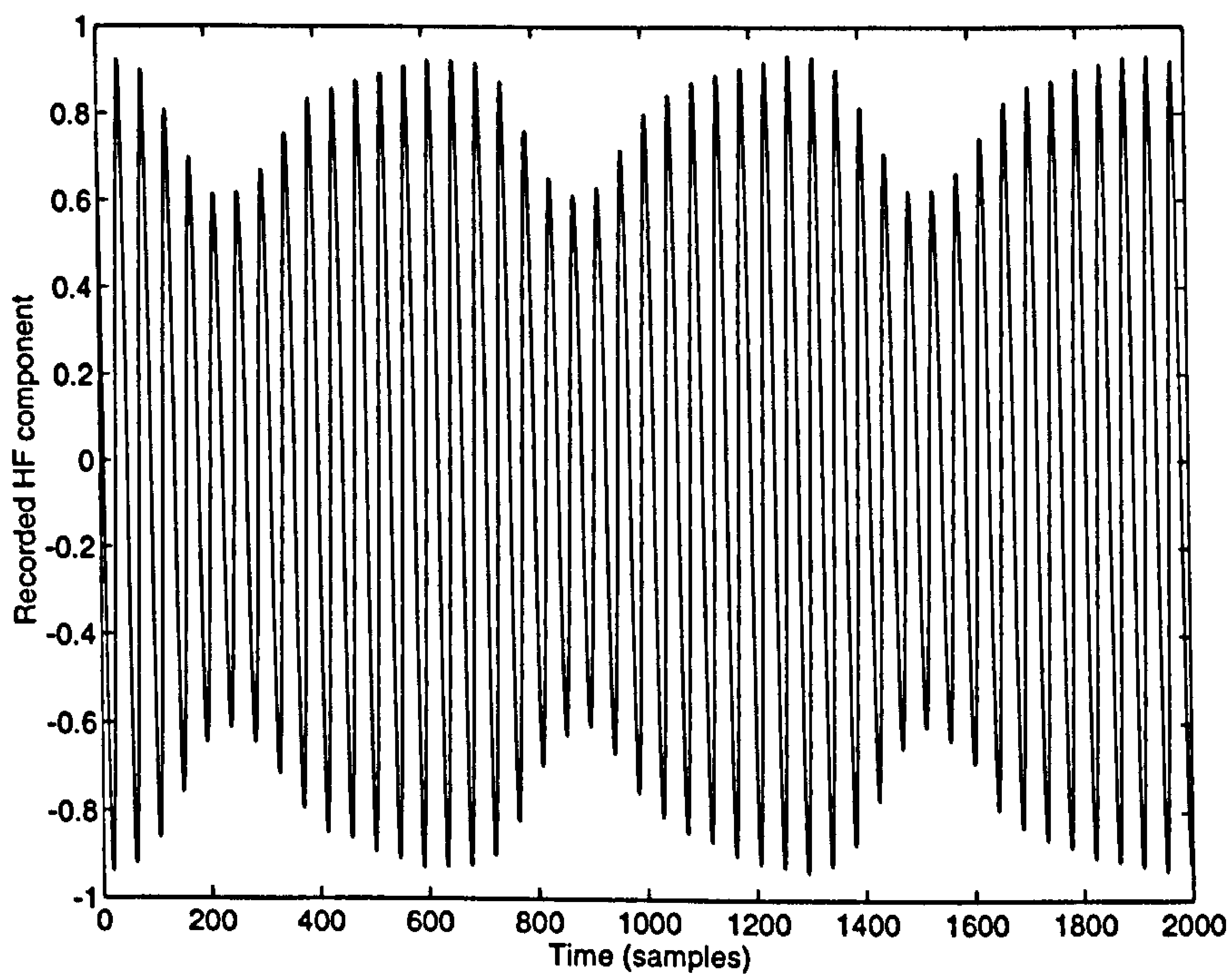
[12] describes briefly a method for measuring the Doppler type distortion using traditional equipment. The method makes use of two sine wave generators to provide the simplest composite signal and uses an FM discriminator to measure the depth of FM modulation. These pieces of equipment were not available. After consideration it was decided to attempt something similar but entirely in the digital domain. The equipment set-up diagram is reproduced below, note that the presentation has been altered to maintain consistency.

Figure 6.12: Equipment for Measuring Doppler Distortion



A very simple detector may be made by subtracting from the signal received the signal transmitted (taking account of any delays), provided that the amplitudes are the same and there is no other form of distortion present this yields good results. If the received signal is taken and the LF component filtered out, as in the graph below, it can be seen that the signal is subjected to considerable AM type distortion - approximately 35%, which is more than an order of magnitude larger than the FM type distortion expected. (It is noted that a 1% change in phase will result in different results to a 1% change in amplitude but for these purposes the simplistic approach is held to be reasonable.) Obviously some other method will be required.

Figure 6.13: The recorded HF component

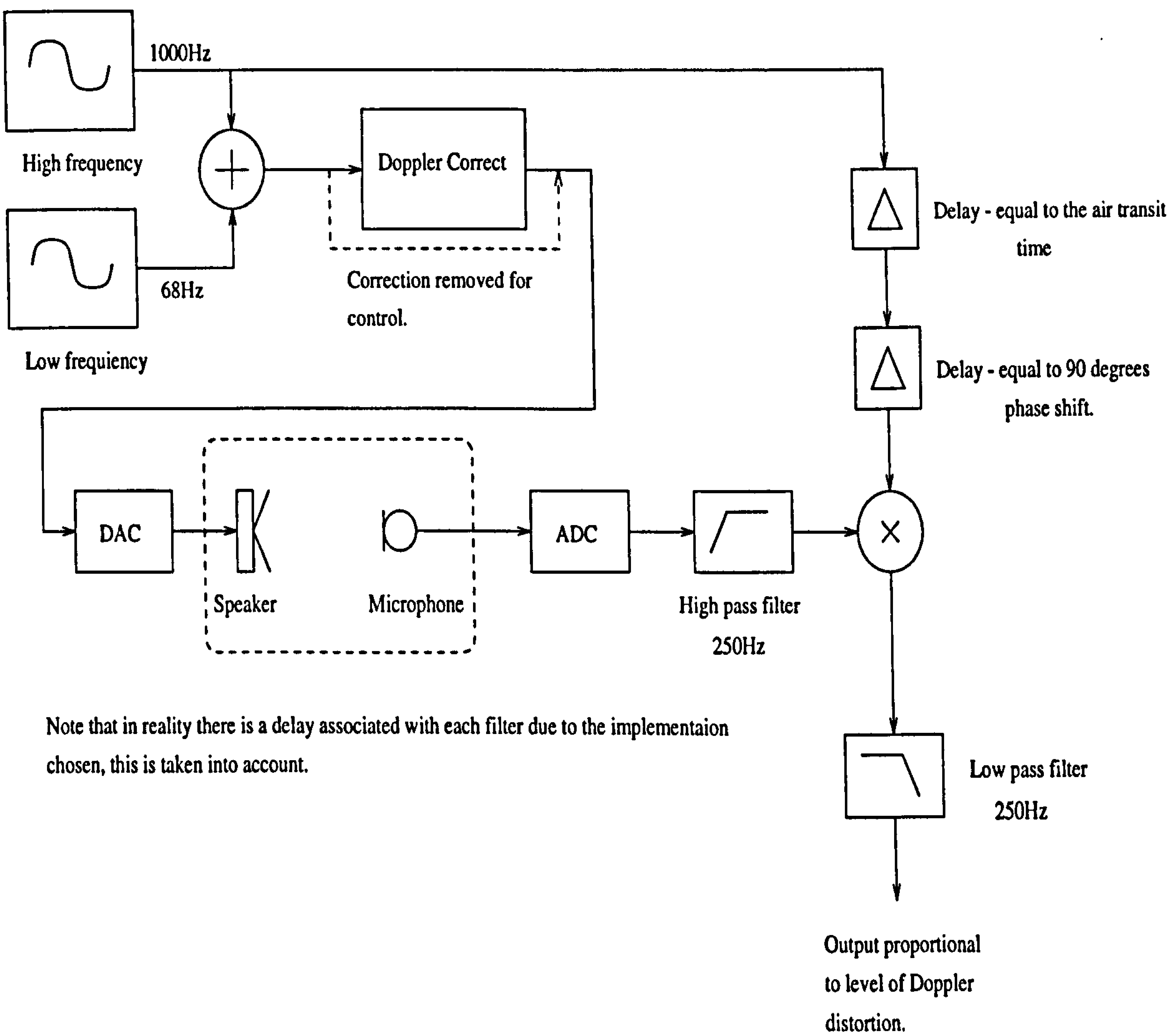


Some difficulty was experienced finding a suitable FM discriminator to implement in Matlab that would be tolerant to addition of noise and absolutely reject any AM signal present. The solution chosen used a DSB-SC decoder, as described in section 6.3.4, and made use of this detector's extreme sensitivity to the accuracy of the local carrier.

Using the analysis presented in the subject review provided that A_c and A_m are constant then this DSB-SC demodulator will provide a very effective phase detector comparing the phase of the carrier and the carrier replica.

Now if this concept is applied to the measurement problem in hand then the "carrier" signal is the high frequency signal and the "message" the lower - the one that causes the majority of the velocity signal. If the higher frequency signal is Doppler distorted by the lower frequency signal then this will show up as FM (or PM) modulation of this signal according to the form of the lower frequency signal. A block diagram for the equipment is given in the figure 6.14.

Figure 6.14: Block Diagram for Doppler Testing Equipment



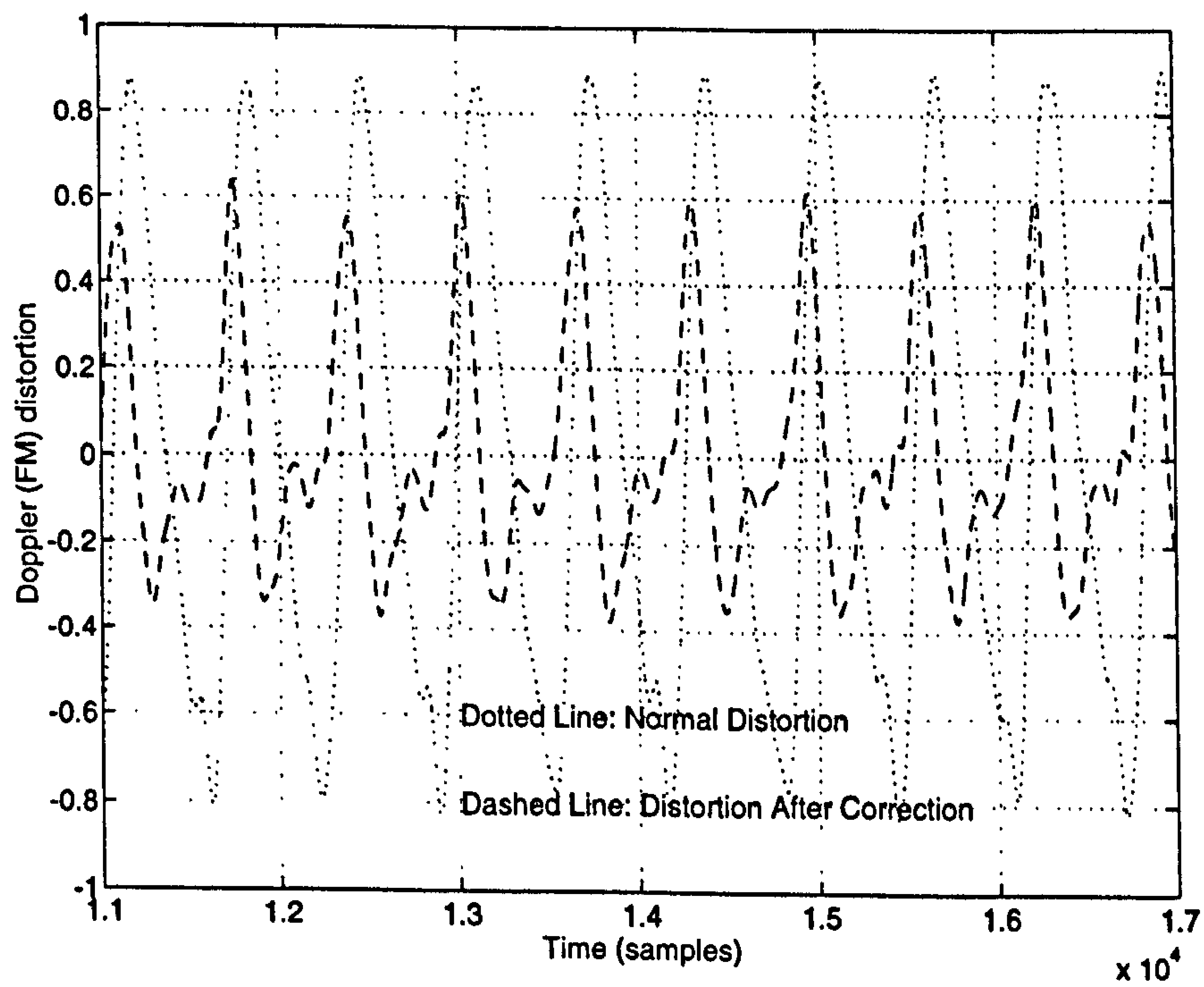
A real speaker will produce some degree of AM distortion, this will cause the value of A_m to vary slightly. Thus the only effect of an AM signal will be to cause a scaling effect to the measurements of the FM component - provided that the AM effect is relatively small the damage will not be significant, especially for comparative measurements. If the desired signal were to have some component at low frequency added to it then this would become an DSB-SC signal and would be removed by the filtering process.

The concept was tested with various distorted FM signals and the results were found to be in good agreement with that predicted by the above analysis.

When this technique was applied to measuring the Doppler distortion level improvement on a real loudspeaker the results were very encouraging. The result of applying this technique is given in figure 6.15. It can be clearly seen that the Doppler distortion measured after the correction is reduced by almost a factor of two. Since these results are based on a real test then it is not surprising that the results achieved are less than perfect correction, especially when the velocity estimate was producing using filter whose values were largely guesswork. It is expected that better results would be achieved with a better velocity estimate.

Given that the sample frequency for the test was about 48kHz and the LF signal was a sine wave at nominally 68Hz then it can be seen from figure 6.15 that the level of frequency domain distortion changes with the LF signal. Referring to section 6.3.4 it can be seen that the measure of Doppler distortion produced by the experimental apparatus is approximately cos(phase error).

Figure 6.15: Measured Doppler Distortion With and Without Correction



6.5 Discussion of Results

It has been established that all speakers will suffer from "Doppler" type distortion, or as it is also know, moving boundary condition effects. Distortion of this type will be at quite a low level since the the cone velocities experienced are typically much lower than the speed of sound. This type of distortion is caused when a loudspeaker cone is called upon to reproduce two signal components simultaneously. The cone velocity at any time will be the sum of the velocities expected due to each signal, provided that the system is linear. This can be considered as a moving source, the result of one signal reproducing the other signal. It is well know that when sound is emitted from a moving source then to a fixed observer the apparent frequency will be altered, the phenomenon most often know as the Doppler effect.

Doppler type distortion is a Frequency Modulation (FM) effect. Comparison of the level of distortion expected with a table of the minimum detectable musical intervals indicates that the levels of distortion experienced should be detectable at mid-band frequencies. The peak levels expected for a representative speaker driven hard are 4 times the audible threshold.

A paper has been published recently [66] that considered the Doppler distortion effect in loudspeakers in great detail. The author even proposed the form that a Doppler compensation device should take. Unfortunately it has proved difficult to compare the proposition due to the very different starting point and methods. The author seems to be working towards a system whose underlying function will be similar to that described here, though he considered the problem in the continuous domain.

A compensator has been developed that should yield very good results and is easy to apply to the digital signals that are common to modern HiFi systems. Simulating the system indicated that the initial concept is based on sound reasoning. In a practical test the system demonstrated a worthwhile improvement - something like a 50% reduction in the levels of this type of distortion even given the limitations of the equipment available. All the engineering constraints to applying the compensator in practice have been considered and there are no difficult problems that remain to be solved.

There have not been any other complete systems reported and so comparisons of function or efficiency are not applicable.

The next stage in developing this concept would be to implement the system on DSP hardware and set up a subjective test.

Chapter 7

Room Acoustic Control

7.1 Introduction

7.1.1 Overview of the Chapter

The aim of the work reported in this chapter is to further develop an adaptive room acoustic compensation system to reduce the colouration introduced by the room during audio reproduction. The target system will use only the existing loudspeakers and will operate by producing sounds in addition to the reproduced music to cancel the effects of acoustic reflections for the listener. The system will require no user intervention during initialization and will adaptively compensate for changes in the environment.

Systems designed to compensate for room acoustic imperfections are not new. Implementations of these systems became practical with the advent of dedicated DSP processors. The simplest method involves measuring the room impulse response, designing a filter and then filtering the signal before reproduction. The only aspect of this that needs to be carried out in real time is filtering the signal, the remainder of the processing can be carried out off-line. SigTech of Cambridge, Massachusetts produce a system of this type targeted at professional audio applications (sound reinforcement, studio monitoring etc.) SigTech's system requires the services of a specialist engineer and additional equipment to set up the filter which will remain constant regardless of changes in the environment.

This chapter looks at developing an adaptive filter system, a system that requires no user intervention when initially setting up and can change with a changing listening environment. The changes need to take place in real time and should be transparent to the user. The system being researched is targeted specifically at the audiophile grade stereo reproduction market and the research will concentrate on the practical problems that would be encountered. An adaptive room acoustic correction system was proposed by Nelson *et al* [44], [46], [45] and by Yasukawa *et al* [64]. Nelson designed a system based on an adaptive filtering concept by Widrow [62] called the "filtered x" algorithm. Nelson also filed a patent for a system that using the same ideas [43]. The development of the current state of the art is considered in more detail in section 7.2. Since the basic technology has already been developed the emphasis of this research will be on the implementation, limitations and possible improvements to the current technology.

The chapter is structured with a review of significant research in section 7.2 following this introductory section. Here, analysis of current research indicates that while the technology involved in fixed equalizers is largely mature, there is scope for research into the problems that may be encountered when implementing an adaptive system for a commercial market. At the start of this section the device marketed by SigTech is examined.

In section 7.3 a target specification of the desired system is proposed. This will be used later to establish the viability of the methods proposed. In brief it should be possible to add the proposed system to existing stereo reproduction apparatus. The system should, in a reasonable amount of time, improve the quality of the reproduction at the listening point.

Section 7.5 deals with the relative merits of the inverse filter systems that have been proposed as well as considering options originating from more general research. There are significant disadvantages to some of the methods that have been proposed that become particularly evident when practical implementation of the system is considered.

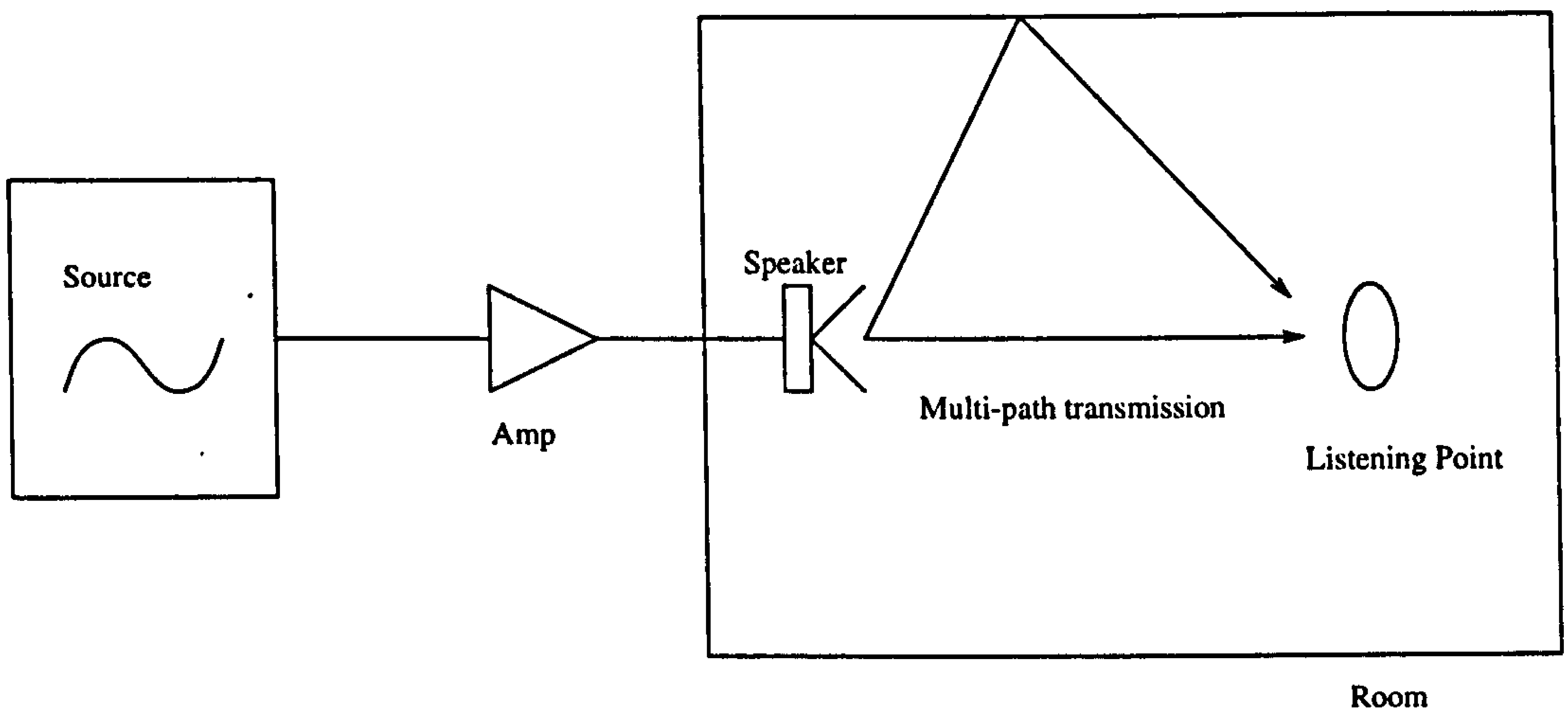
Having established which of the possible architectures appears the most promising further possible improvements are considered in section 7.6.

Finally, the limitations that a real, commercial, system would face are considered in section 7.7. The results of this analysis indicates that there are significant and apparently insurmountable limitations that will severely limit the possible operational bandwidth of the proposed system.

7.1.2 Overview of Coherent Acoustic Equalization

When a sound is reproduced in a real room there will be many reflections from the walls and contents of the room. The listener will not only hear the desired sound, that direct from the loudspeaker, but also many slightly late reflected versions. The reflections are very much a function of the size and shape of the individual room and furnishings and thus are not easily predictable. Figure 7.1 shows the problem pictorially for a single reproduction source.

Figure 7.1: The Room System



The effect of the room acoustics is normally to "colour" the listener's perception of the sound. (Reflections will not be heard as distinct signals provided that the delay is not more than about $50ms$ [22] *pp 440*, which is likely since this represents a distance of $17m$.) Probably the most significant corruption to the acoustic illusion will be that an incorrect impression of image size will be given, since the recording will probably have been made in a much larger room. Acoustic information encapsulated in the reverberation present in the recording will be confused by the additional echoes, with probably very short delays, due to the listening room. (As noted before, in section 3, real recordings tend to be constructed from a large number of close microphone recordings; these will have been given the correct ambiance using electronic effects.) The concept is to generate a filter that will remove these reflected sounds, at least as far as the listener is concerned.

There are two modes in which an acoustic correction filter could operate:

- Absorption Mode

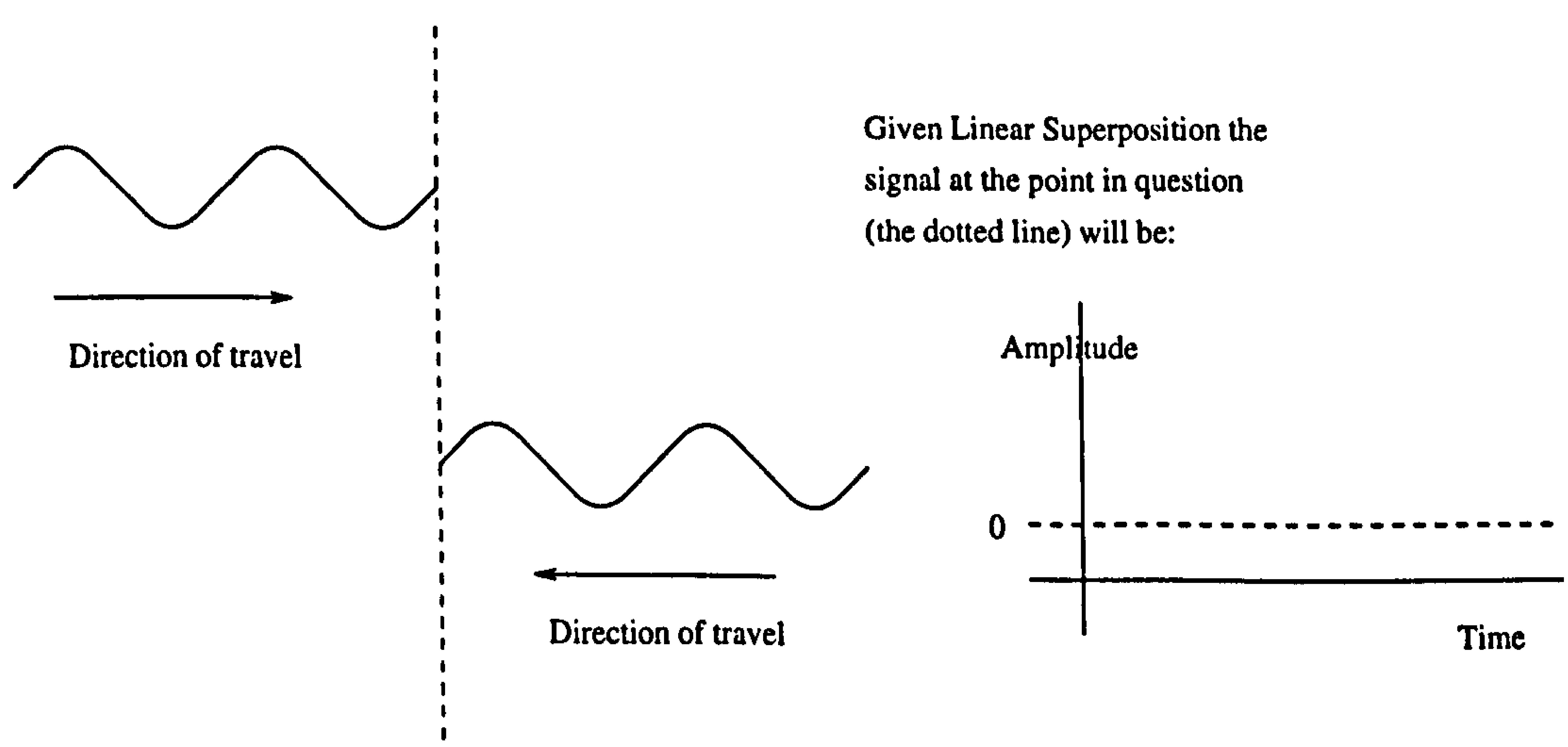
  The filter system could be designed to absorb all of the sound energy except for that passing directly from the loudspeakers to the listener. This method is difficult to engineer without a very large number of correction sources, ie. "speaker walls" [27]. The technology required to achieve reasonable results with this type of system, however, is not of a particularly high level. Systems of this type have been demonstrated for many years, though not generally for the HiFi application, [47] being a late example.

- Destructive Interference

The alternative is not to attempt to remove the sound energy from the air, only locally cancel its effects. This method relies on an effect known as linear superposition. (The compressibility of air may be considered linear to normal audio level signals.) This latter method is much more interesting since it is possible that reasonable levels of correction could be achieved using only the existing two stereo speakers as sound sources. Unfortunately, the level of technology and complexity of the filters required to achieve this aim is considerable. Overall this is the only method that will be considered in further detail here.

The operation of destructive interference is not difficult to understand. Signals passing through a medium, such as audio signals through air, have direction, amplitude and age (time since they were emitted). For sine waves the age may be expressed as a phase at the point of interest. If more than one signal passes through a point the signal perceived at this point over time will be related to all the signals passing through. In the general case the signals may affect each other as they pass through a point - the observed value at the point will certainly be a function of all signals. With audio signals travelling in air it is reasonable to assume that the air's response will be linear and thus the signals will not affect each other and the response at a point will simply be the sum of the two parts. An example of destructive interference is given in figure 7.2.

Figure 7.2: Disruptive Interference



Thus, for audio reproduction, if the signal at the listening point is the linear sum of the desired signal and reflections from the walls then the effects of these reflections can be cancelled by an additional signal. The additional signal needs to be exactly equal and opposite to the sum of the reflected signals at the listening point when it arrives at the listening point. (The correction signal will also need to cancel the reflections at the listening point of earlier correction signal.) This means that the correction signal needs to be made before the reflected signals that it is to correct have arrived at the listening point. Since there are a very large number of reflection paths predicting the reflected signal is difficult.

Provided that a suitable correction signal can be calculated then over time the net result is that the desired signal is reproduced correctly at the listening point without any corruption from reflections.

There have been graphic equalizers marketed that claim to compensate for the room's acoustic by adjusting the relative levels of the frequency bands. If the listening room's steady state frequency response is measured then it is possible to provide steady state correction by this method. Simple equalization of this sort will not work well for audio reproduction due to the transient nature of audio signal; the room's steady state frequency response is not useful when evaluating how it will respond to a transient signal unless absolute phase is preserved. By way of example consider a continuous sine wave reproduced in a reverberant room; the signal could be corrected at the listening point by adjusting the amplitude and phase. Now consider a similar but pulsed sine wave. Simply changing the amplitude or phase of the sine wave will not correct for the effects of the reverberation in this case. In addition the complexity of the room acoustic frequency response presents a problem, there are something like 175,000 room modes in an office-sized room below 3.5kHz [26].

Generating the signal to destructively cancel the effects of the room acoustics is not a trivial matter. Nothing can be assumed about the listening room's dimensions or contents - though it is far from certain that any approximate information would be useful due to the number and complexity of the reflection paths. A realistic possibility is to place a microphone at the listening point and then make use of adaptive filter generation technology to generate a correction filter. The filter would then be fed with the desired signal and would generate a correction signal to be added to that sent to the loudspeakers. Thus the desired signal is perceived un-tainted at the listening point.

The problem is further complicated when it is considered that a stereo system has two channels and we have two ears but there is nothing to prevent crosstalk. Thus some of this signal originating from the left speaker finds its way to the right ear and vice versa. It is possible to attempt to control these signal paths too, as was suggested by Nelson *et al* [44].

The problem is deciding what the optimum response for these crosstalk paths should be. (The crosstalk paths are g_{12} and g_{21} in figure 7.3 and the filters to control these paths are h_{12} and h_{21}.)

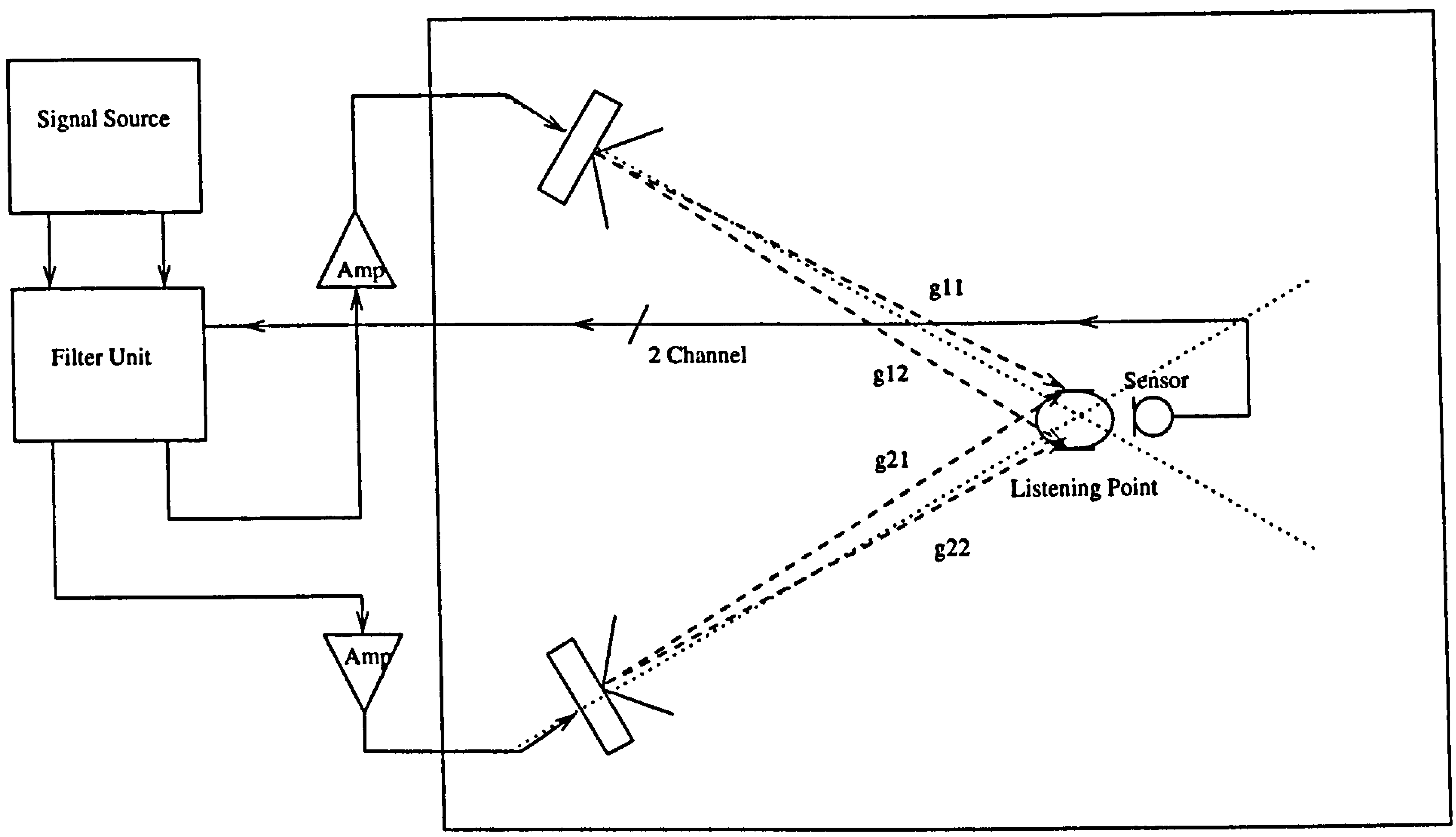
There are many other practical problems that have to be considered and limitations that have to be assessed before a practical system could be designed.

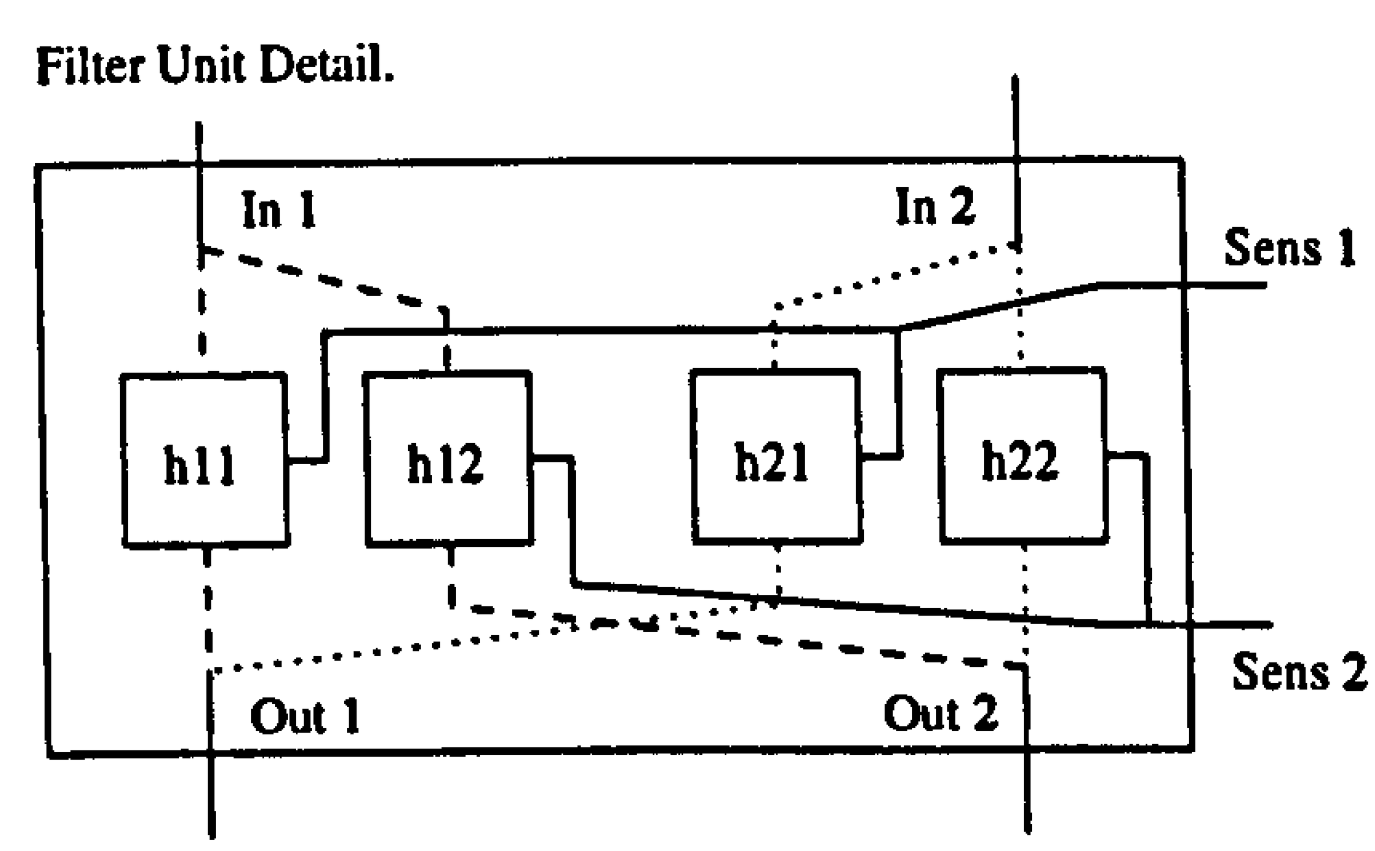
7.1.3 The System

The concept is to generate a signal that will destructively interfere with all reflected signal and thus leave only the desired signal at the listening point. In order to achieve this it will be necessary to insert a digital filter between the amplifier and the source, perhaps in the digital output line of the CD player. The filter will need to be re-calculated for each room and for any change within the listening room thus adaptive methods should be used for the calculation. Adaptive systems need a performance measure which means that there will have to be a sensor at the listening point.

The system proposed by Nelson *et al* and others is show in figure 7.3. This figure also shows the various acoustic transmission channels and indicates the symbols that will be used throughout this chapter.

Figure 7.3: The Inverse Filter in a Stereo System.





g11 = Room impulse response from speaker 1 to ear 1
g12 = Room impulse response from speaker 1 to ear 2
etc.
h11 = Correction filter for signals from speaker 1 to ear 1
h12 = Correction filter for signals from speaker 1 to ear 2
etc.

The signal from the source is filtered by a correction filter before being fed into the loudspeaker and therefore the room. The filter, say h11, alters the signal so that at the same time as a reflection from an eariler portion of the signal reaches the listener an exactly opposite version will reach the listener from the loudspeaker. This means that the listener will never hear any of the reflections from the walls. The paths g11 etc represent the impulse response of the room, that is they are the effect the room will have on a signal from the apropriate loudspeaker but observed at its destination.

7.1.4 Room Modes and Other Acoustic Effects

Traditionally a study of room acoustics has split the effects that have been observed due to room boundary reflections in to a number of different categories. The effects are all caused by the same mechanism, superposition of reflected sounds, but are subjectively quite different. The following categories are typical:

- Standing Waves
- Room Modes
- Flutter Echoes and Other Unusual Reverberation Effects

Standing Waves

When a room is excited with a continuous sine wave then standing waves will result from the superposition of the direct and reflected signal. In places the waves will superimpose additively and result in a higher sound pressure than might have otherwise been expected, and similarly in other places the superposition will be subtractive. The locations where the superposition is additive are typically called nodes and the locations where it is subtractive, anti-nodes.

Steady state models of room performance are not generally useful since audio signals are transient in nature and thus a gain/relative phase correction will not be sufficient to deal with the (inherently) delayed reflections.

Room Modes

Room modes are the resonant frequencies of the room and thus the average amplitude of the signal will tend to differ markedly from the source amplitude. (In addition to the location specific effects caused by standing waves.) Superposition will always occur, but will only build into strong resonances at the room's resonant frequencies - the room modes.

For a continuously excited system the effects of room modes can be significant, the measured sound pressure can be several times the input pressure [14]. The frequencies at which these maximums occur is given by Raleigh's equation [56]

$$f = \frac{c}{2}\sqrt{\left[\frac{n_l}{l}\right]^2 + \left[\frac{n_w}{w}\right]^2 + \left[\frac{n_h}{h}\right]^2} \qquad n_l, n_w, n_h \quad integer \qquad (7.1)$$

The integers n_l, n_w and n_h represent the room mode number for the length, width and height respectively.

Raleigh's equation is an expression which gives the frequencies of the axial, tangential and oblique modes of a room was derived by Raleigh by solving the wave equation for a plane wave in three dimensions and imposing the necessary conditions at the room boundaries.

It is possible to compensate for modal effects at a single point within a room for constant signals by using normal (gain and relative phase) frequency domain compensation. This compensation will not be valid for normal acoustic signals since they are transient in nature and thus attempts to use graphic equalizers to make the reproduction more accurate subjectively does more harm than good.

Flutter Echoes

Flutter echoes occur if two opposite walls in a room are significantly more reflective than the other surfaces. This effect occurs since the echoes from these two surfaces are dominant and, if the distance is sufficient, this dominance allows the human auditory system to separately identify these pulses of sound energy.

Room modes and Flutter echoes are both prime targets for any acoustic correction process since these are strong, noticeable effects that significantly reduce the subjective performance of the room. Room modes start occurring at low frequencies, even for normal sized rooms. For instance, taking a room of 5*4*2m the lowest resonant frequency (mode) will be at 34Hz and there will be modes at 42.5, 54.4, 68.0, 80.1, 85, 91.5, 95.0 etc. (With something like 65 modes below 200Hz.) Square (or worse yet cubic) rooms will suffer from very strong room modes because of the "repeated" resonances.

The effects of room modes (generally and specifically to a location) can be predicted mathematically but this would require accurate measurements and a very well defined room. In practice it is much easier to make use of some automated system, such as the adaptive filter suggested in this chapter, to correct the signal.

7.1.5 Experimental Methods

Pseudo Random Binary Sequency (PRBS) methods were used to measure the room impulse responses. Sequences of this type allow accurate impulse response measurements even in the presence of noise. Full details are given in Appendix 9.1.

Two principal methods were used to calculate the inverse filters. The "Matrix Method," which is described initially in Appendix 9.1, is an expression of the problem that allowed the use one of Matlab's built in approximation algorithms. The techniques used in Matlab could not be used for a real implementation but the answer obtained represents the realistic optimum results that could be obtained by practical adaptive algorithms.

The generic adaptive algorithm that is used for adaptive inverse filtering is also described in Appendix 9.1.

7.2 Review of Significant Research

The concept of using a filter to correct for room acoustics is not new, but the application of real-time adaptive technology is relatively new. Recently there has been a similar development of DSP processing power to the growth of general microprocessors and this new power is opening up more complicated problems.

7.2.1 Commercial Devices

Commercial systems based on fixed filters are available that are able to compensate for the effects of reflections in the listening room, one such device is manufactured by SigTech. This device operates in two stages, there is an initial measurement/filter design stage and then the filter remains constant during operation. Information concerning the SigTech device was obtained from their WWW site [57].

The principal of operation is described by Genereux in two separate conference papers, [23] and [24].

In these papers Genereux describes a fixed equalizer system designed to compensate for both the room acoustics of the reproduction room and linear response imperfections in the reproducing loudspeakers. He notes that the area over which the equalization is subjectively useful is a problem even for fixed systems. (Fixed systems do not suffer from the sensor having to be a small distance from the listening point and thus the bandwidth is not inherently limited.) The system makes use of a "segmented filter", which is also the subject of a US patent application, to reduce this problem.

The SigTech system only attempts to correct for "early reflections", those arriving with 50ms of the direct sound, as the author claims these are the most damaging to the listener's sense of tonality. This limitation reduced the size of the filter required dramatically and thus reduces the scale of the problem.

Genereux suggests that to reduce the sensitivity of the system performance to listener location that the frequency resolution of the filter should be high at the extreme LF, but should reduce with frequency. This should reduce the effects of highly location specific occur if the frequency resolution were left constant. (In practice it will tend to deactivate correction for room acoustic effects at the higher frequencies but may still provide useful correction for deficiencies in the loudspeakers.) In order to achieve this Genereux devised what he called the "segmented filter."

The operation of the "segmented filter" is described in the conference papers [23] and [24]. The exact method used to calculate the segmented filter is not given but from what is said the probable process has been extrapolated.

The "segmented" filter used is a nominal 2450 taps in length, which is a little over 50ms at a sample frequency of 44.1kHz. This is divided into 8 segments in order to control the time-frequency resolution. Each segment's bandwidth is one octave greater than the previous and the number of coefficients dictates the frequency resolution. (Roughly this is a wavelet method given this limitations of sampled time.)

- Measure impulse response

  The impulse response of the room is measure at a small number of locations around the listening area. Responses that show strong room acoustic effects or those that are markedly different from other measurements in the listening area are discarded. The listening area is approximately $1m^2$

- Calculate the HF filter

  The HF filter is calculated directly from the impulse response. An adaptive method is used to give an approximate filter. (Exact methods for calculating the inverse of a finite impulse response could have been used, but these are only exact compared to the finite (truncated) impulse response they were calculated against.) The length of all of the filters before interpolation is the same at approximately 20 coefficients.

- Calculate the next filter

  The impulse response is then decimated by a factor of two with appropriate filtering to prevent aliasing effects. Another filter of the same length as the previous one is estimated. This filter is then interpolated by a factor of two.

- Calculate the next filter

  The impulse response is decimated by a further factor of two with the appropriate filters to prevent aliasing. Again a filter of the same length is calculated but this time it is interpolated by a factor of four.

- Repeat

The process is repeated until all 8 filter segments have been calculated.

- Calculate final filter

  The filter segments are then added together using scaling factors that may be user determined.

The scaling factors are typically in the rage 0.1-0.4 and are applied to each of the segments. Setting the scaling factor to 0.0 does not disable the correction in that segment, but limits the correction to the resolution of the next highest segment. One way to prevent correction above a certain frequency is to set all higher segments to zero

Thus if the longest filter (after interpolation) were used alone it would deal only with the LF elements of the room transfer function. This means that the HF filter would be (in time) 2^7 shorter than the LF filter - a factor of 128. Thus the HF filter would be 0.5ms long. A filter of this length could have no useful application in room acoustics correction. The path length the filters are equivalent to would be 19m, 9m, 5m, 2m, 1m, 0.5m, 0.3m and 0.15m. Since the realistic minimum acoustic path in a normal room will be at the very least 5m the last filter to have any acoustic correction ability will be operating at an effective sample frequency of something like 700Hz, ie a bandwidth of 350Hz. The higher bandwidth filters would be able to improve the response of the loudspeakers only.

The short filters used in the SigTech device would not be able to meet the PMR and PPMR targets (see section 7.4) that have been set for the adaptive system. Short filters tend to be unacceptably "noisy" - they could not be good enough at the lengths suggested here even if the two stereo channels where used separately to allow more effective filter power.

The SigTech device is aimed at the sound reinforcement/PA environment thus it is essential that the delay introduced by the device is slight. A modelling delay of 17 ms is employed. For the domestic market this limitation need not apply, which allows the use of longer delays.

The SigTech unit is different from the concept under investigation since it does not allow real time adaption and correction and thus has no mechanism for tracking slowly changing parameters. Alterations to the listening room require that the equalizer is re-setup, rather than the device simply adapting to its new environment as the concept being researched here would.

7.2.2 Published Research into Adaptive Inverse Filters

Considering the developments in chronological order.

In 1985 Widrow *et al* published a book titled "Adaptive Signal Processing" [62] which contained a complete review of Adaptive techniques and an entire chapter discussing "Inverse Adaptive Modelling". (The problem of finding inverse filters to compensate for the room acoustics can be classified as inverse adaptive modelling.) Widrow *et al* take the minimum inverse filter system and develop it to produce the "filtered x" algorithm and beyond.

Widrow *et al* do not suggest compensating for room acoustics as a possible application for this technology but this omission is predictable since, at the time the book was published, the technology was not readily available to implement such a system. The example applications given in the book are typically for much simpler or much slower systems.

In 1988 a paper was published titled "Inverse Filtering of Room Acoustics" [39]. In this paper Miyoshi *et al* consider first the basic application of an adaptive inverse filter system and then extend this over multiple channels in order to give a "greatly superior" solution. The multiple channel inverse filter system is proposed as a solution to dealing with the problem that the room transfer function is non-minimum phase.

In a section of the paper titled "Review of Conventional Inverse-Filtering Method" the problem of creating an inverse filter is expressed in such a way that it will be unsolvable by LSE methods as a justification for using multiple channels. Adding channels to the problem when expressed this way will not allow the creation of a inverse filter for general room acoustic problems. (The only true solution is to allow the filter to be non-causal, either by including a delay or by delaying the required response. See Appendix 9.2.) Provided that the problem is expressed in such a way that a stable filter with converging coefficients is produced then adding extra channels will improved the performance of the filter.

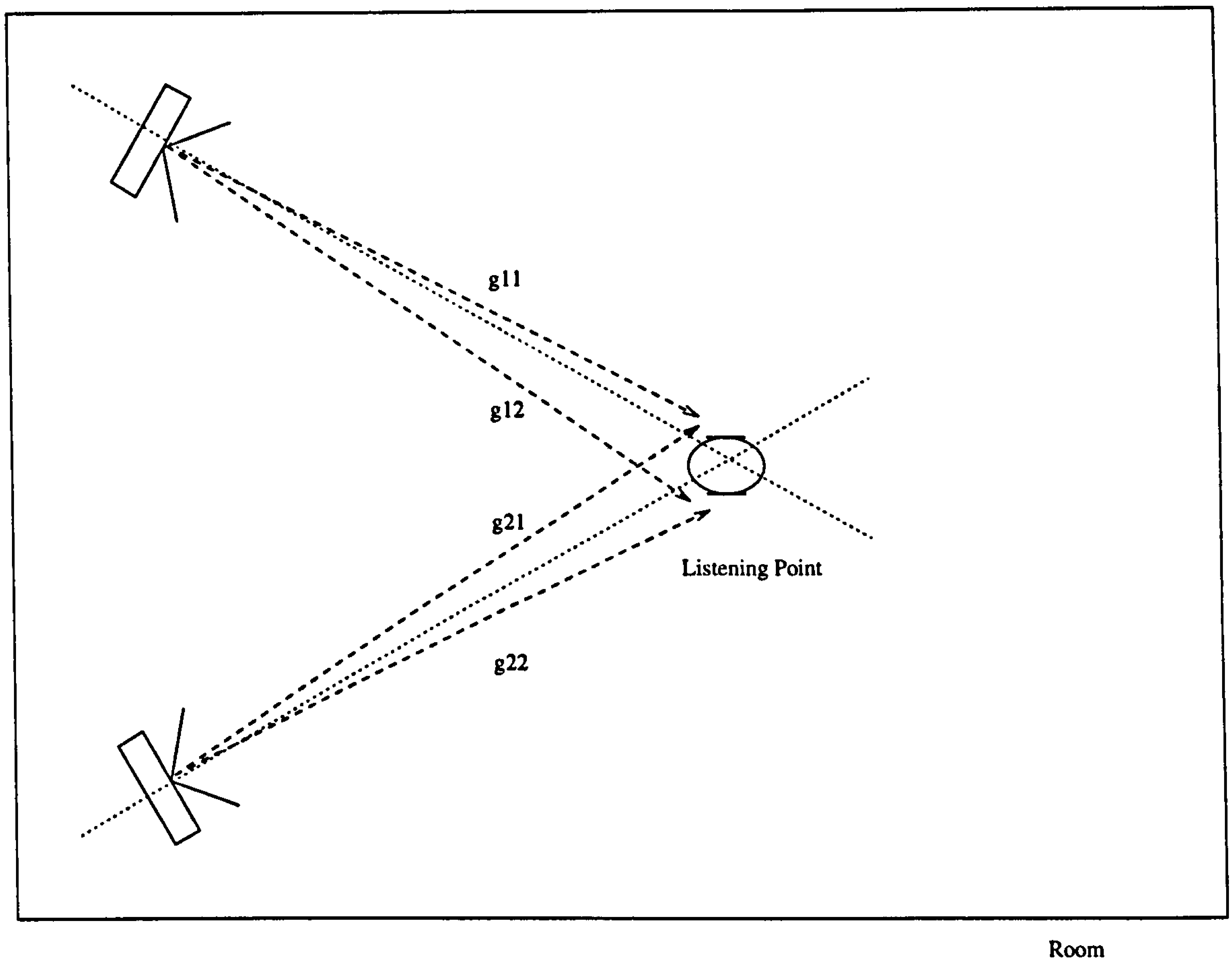
There is a finite transit time associated with the system thus the inverse filter must be allowed to be non-causal to at least this degree. This can be easily accommodated by adding a delay. Another way of looking at it is that if the room impulse response is zero at time zero on all channels, which is highly probable, then there is no value that any causal element of the inverse filters can take that will yield a non-zero result at time zero in the system impulse response. FIR filtering is equivalent to convolution. Thus, the only values to affect the time zero system response are the first value of the room impulse responses and the first value of the filters. Since the first value of the room impulse responses is zero the result will be zero. (The room impulse response is a non-minimum phase system since there are coefficients in the response that are larger than the first.)

Miyoshi *et al* also state that it is not possible to get the error power of a single filter to converge to zero regardless of the filter length for a single channel using an LMS filter generation technique. This represents a serious problem if true for acoustic correction systems. Significantly, however, the desired response they use is always taken as an impulse at time 0, which can easily be demonstrated as an impossible target for any causal inverse filter system. In appendix 9.2 it is demonstrated that the error will converge for a single filter provided that it may be non-causal, thus it is not necessary to employ multiple channels in order to achieve good filter performance.

Subsequent to the publication of [39] a patent application was filed by Nelson *et al* [43] titled "Improvements in or Relating to Sound Reproduction Systems."

In this patent applications Nelson *et al* describe a complicated inverse filter arrangement for application to stereo reproduction. In this system not only is each channel optimized to its respective ear, the signal path from the other speaker is also controlled. The signal paths are shown in figure 7.4.

Figure 7.4: System Proposed by Nelson *et al*



It is suggested that paths g_{12} and g_{21} are minimized while paths g_{11} and g_{22} are optimized to suit the application; for reproduction that is most accurate the latter paths are expected to be optimized towards an impulse. (Anything convolved, i.e. filtered, with an impulse remains unchanged. Thus this is the ideal target for a perfectly uncoloured reproduction system.)

The application is not clear on exactly the methods chosen to produce the inverse filter and gives no consideration to the limitations of the proposed system. It is interesting to note that a "modelling delay", the delay required so that the inversion problem becomes tractable has been included explicitly by Nelson *et al.*

In 1992 Nelson *et al* published two papers in different journals [44] and [45] which both gave more detail of their proposed inverse filtering method. The layout of the system in [44] is identical to that in the patent application but the concept has been extended to many channels for [45]

Considering [44] initially. Some of the problems of applying this system to stereo signals are hinted at early in the paper when Nelson *et al* use artificial head binaural recordings as an example of a two channel stereophonic system. The system described exists, but almost all current recordings are designed for the "stereo" reproduction system first suggested by Blumlein [54]. (The origins of the stereo reproduction system and the mechanism by which it works are examined in some detail in section 3.) Later Nelson *et al* state that "...the approach has been generalized such that it can be applied to any recording technique involving sensing the sound field with any number of transducers. It is thus not only applicable to artificial head or coincident microphone techniques [presumably the stereo system is included here] but also more recent approaches..." Unfortunately they fail to give details of the exact method by which they hope to apply their technique.

Nelson cites two previous papers [19] and [11] as having demonstrated single channel inverse room acoustic filters.

Nelson *et al* state "However, we cannot compute an exact inverse for the matrix $\mathbf{C}(z)$ [the room transfer function] due to the presence of non-minimum phase components." As already discussed this is misleading. It is true that causal inverse filters do not exist for most acoustic inverse filtering problems but provided that the filter is allowed to be non-causal then the coefficients of the filter will converge, see appendix 9.2. For that reason any finite acoustic inverse filter will be an approximation, but as the number of coefficients becomes large then the error will tend to zero. The presence of non-minimum phase components means only that the filter cannot be causal. An exact filter will generally extend in both directions to infinity, but the coefficients will also converge in both directions. A non-causal filter can be implemented in a real system if a delay is allowed.

The LMS method chosen by Nelson *et al* is Widrow's "filtered x" structure generalized to operate for multiple channels [62]. Nelson *et al* correctly note that this choice requires a model of the room impulse response that is never more than 90^o out of phase with the real room if the method is to converge. (Any inaccuracy in the model will reduce the rate at which the filter will converge.) The alternative solutions are not considered and no reasons are given for the choice of this method. Widrow *et al* state that the filtered x algorithm will converge more accurately in the presence of noise and is thus likely to be preferable in most situations. Widrow suggests more advanced forms of the filtered x algorithm that do not need a model of the room impulse response but Nelson *et al* do not appear to have considered these options.

The filters used by Nelson *et al* in their practical work were extremely short, due to the limited processing power available to them. The room models used had only 64 taps at a sample rate of 6.4kHz. (This corresponds to a filter of length of 0.01 sec or a maximum path length of 3.4 meters. The experimental room used was approximately $25m^2$, assuming this to be a cube gives sides of length approximately 3 metres. Thus, only the very shortest reflection paths could possibly be adequately simulated by the room model used and therefore it seem unlikely that the experiment would be able to produce useful results.

Of the results given the frequency responses are the most convincing. Unfortunately Nelson *et al* neglect to give the responses *before* adaption and thus these results do not provide any real indication of the system performance. It is also interesting to note that the adaption constant reported is very low, 0.01 to 0.05. The results detailing the area over which the filters are effective must be treated with suspicion since the filters used were too short to deal with any real reverberation from the size of room in which the tests were carried out.

Nelson *et al* have not considered the possible effects of the crosstalk cancellation filters with a real stereo signal. Provided that the signals fed to each channel were orthogonal then the adaption will proceed as expected, but this is not a reasonable assumption for stereo signals. Considering a signal recorded from directly ahead using the coincident crossed cardioid method (or for that matter a dummy head) then it is apparent that the signals to each channel will correlate exactly. The crosstalk cancellation filter will thus see a large quantity of a signal it recognizes (but actually originated from the other channel) and will attempt to suppress it. Since centre and partially centre stage signals are common in stereo signals the only solution would be to use a specially designed training signal to set these filters.

The "error scanning" method, used in the practical demonstrations Nelson *et al* gives, is referred to in [44] and [46]. The concept is that the filters are updated on alternate or fewer samples so that the available processing power can deal with a more difficult problem. Nelson *et al* state that applying this technique did not materially affect the adaption rate. This seems unlikely. It is well known that recursive, or partially recursive (multi-sampled) LMS algorithms converge faster. Thus, adapting on every second sample must reduce the convergence rate.

A much shorter paper by Nelson *et al* [45] covers very similar ground to the previous one but in less detail. The only additional piece of information in this paper that appears of interest here is "...the use of a number of loudspeakers which exceeds the number of points at which sound reproduction is required can have advantages in the performance achieved by the inverse filters." This is predictable since it is possible to conceive of a situation where, when reproducing a single frequency, the listening point happens to be at a node in the standing wave pattern. In this case this frequency could not be produced at the listening point continuously regardless of the level at which the source operates. Adding another channel introduces an additional path to the listening point and thus it becomes less likely for this situation to occur.

In practice the listening point is unlikely to be *exactly* at a node and thus in real systems it will always be possible to equalize the response, but high levels may be required. Since it takes time for the filter response to deviate markedly from its starting point, additional channels that have easier paths to the listening point will improve performance.

In 1994 Chang *et al* produced a paper titled "Inverse Filtering of a Loudspeaker and Room Acoustics using Time-delay Neural Networks" [10] the subjects of linearizing loudspeaker performance defects, including those of a non-linear nature, and compensating for the effects of room acoustics are considered in parallel. The details of the loudspeaker correction are not relevant here. The room acoustic method proposed uses a Time-delay Neural Network with two hidden layers. For the demonstration the network was trained with a simulated room impulse response. (The use of an impulse response as the room model means that the response is inherently linear. This is reasonable since non-linear effects in acoustics are generally held to be very slight at normal audio levels.) The results of applying this technique are not presented in a form that allows comparison with the performance of the more normal NLMS techniques.

It should be noted that FIR filters are a specific and very simple case of a neural network and thus the method proposed by Chang *et al* is not perhaps as complete a departure from the previous work as it first appears.

In 1995 Nelson, Hamanda and Orduna-Bustamante produced another paper with the title "Inverse Filter Design and Equalization Zones in Multichannel Sound Reproduction" [46]. In the early part of the paper, Nelson *et al* examine the conditions under which the room impulse response can be inverted. They conclude that the transmission path to a microphone must not contain common zeros in the transfer functions to all potential sources and that a modelling delay is required such that at least one of the impulse responses of the transmission paths is non-zero. He notes that the latter constraint was omitted by [39], as was also noted earlier in this analysis.

An additional and interesting limiting case is noted. If the earliest (or latest) arrivals to any pair of microphones originate from a single source then inversion will not be possible since the matrix requiring inversion to produce the solution will be singular. In practice this should not occur with a stereo system but it may become a problem if the equipment is badly set up.

Nelson *et al* have plotted the extent of the equalized zones using their multi-channel method. The measure used is related to the error between desired impulse response and that achieved and is partially simulated. (The filters are probably obtained from a real test, as is the new impulse response to the moved microphone, but the effect of the movement is probably simulated using this information rather than being measured directly.) The measure of error used is such that the 0dB contour, the boundary between the area where the filter improves the response and that where it makes it worse, occurs where the mean squared error in the representation is of the same order as the desired response. This measure of accuracy, while convenient to calculate, is of little value when it comes to appraising the performance of the filter. No real indication is given of the relative levels of the approximation to the desired impulse compared to the remainder of the response that can be considered as noise. A more detailed discussion of the measures of filter performance will follow later in this chapter. Regardless the 0dB contour represents a rather considerable error and even the -20dB contour is probably unacceptable in a HiFi environment. The permissible error rate will be considered in more detail later.

Nelson *et al* note that the 0dB error contour roughly coincides with the wavelength of the signal being considered.

In their conclusion Nelson *et al* state that they have demonstrated the requirement that the number of loudspeakers must exceed the number of microphones in accordance with Miyoshi's findings [39]. This is at variance to their statement in the body of the paper where they state that exact inversion is possible, with specific limitations, when the number of sources and microphones are equal. They also state that the number of coefficients required in an inverse filter, if an exact inverse is to be calculated, can be predicted. It will be demonstrated in Appendix 9.2 that this is not generally true, most inversions will require infinite length inverse filters for the filter to be exact. It would be involved to find and demonstrate the exact error, it is easier to simply demonstrate that the result cannot hold generally.

7.3 System Specification

A specification has been drawn up to give some target to aim towards and some indication of methods that are unlikely to ever achieve useful performance. Some aspects of the specification cannot be clearly defined until real subjective tests can be performed.

The system will add directly to existing stereo reproduction equipment without the need for additional reproduction sources. The required sensor will be as unobtrusive as possible. The system operation should be completely automatic and should not require initialization - particularly processes that make use of noise signals.

- "Noise" to be lower than -96dB

  Real noise is not likely to be a problem since all of the additional processing will be carried out in the digital domain, however, the application of the filter will not be entirely beneficial. The problems are best seen in the time domain response. It is difficult to simply describe the imperfections that will occur in the time response, but these need to be carefully controlled. This topic is dealt with in substantially greater detail later in this chapter.

- Frequency response of the system not to be impaired.

- Adaption time to be of the order of 1 minute, but less than 10 minutes.

  The adaption speed needs to be sufficient to track the changes in the room acoustic as they occur. (For the Normalized Least Mean Squares (NLMS) algorithm rapid adaption occurs for about 10*length(filter) iterations of the adaptive cycle, see figure 7.19, and there are further improvements for about 100*length(filter) provided that background noise levels are low. Thus for a 1 second long filter the initial adaption will take 10*length(filter) ie. 10 seconds.)

- Control

  The degree to which the inverse filter acts needs to be controllable. That is it must be possible for the user to vary the level of correction; the SigTech unit provides the user with control over both the ratio by which correction is applied (allowing the user to have half as much correction as the calculated filter would have provided for instance) as well as limiting the maximum correction level. There also needs to be a mechanism to ensure that a filter is not applied to the audio signal until it is sufficiently mature to ensure that the overall sound quality is improved. If a new filter were simply placed in the audio signal path then the resulting sound quality would be subjectively reduced until the filter adaption had progressed.

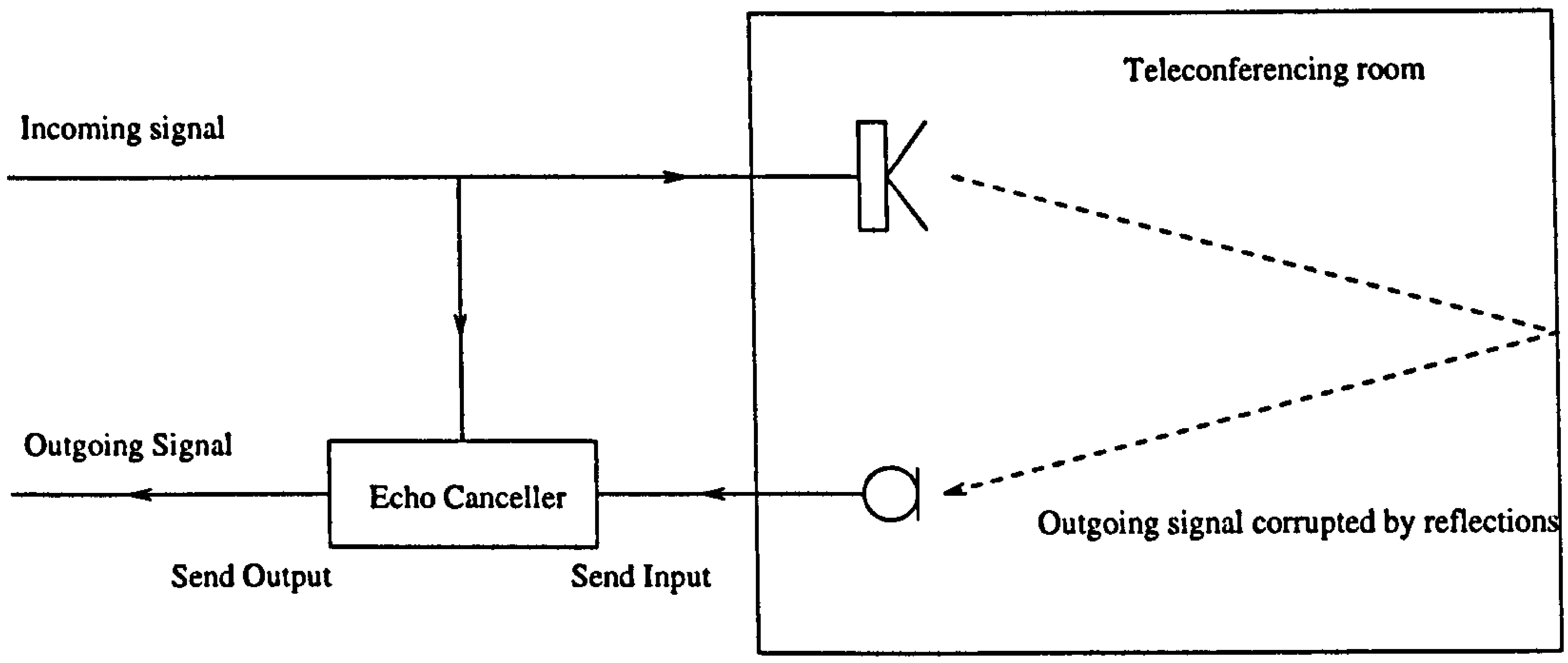
7.4 Measures of Filter Quality

A measure of the quality of the inverse filters produced by the various methods is required. The method must not be tied to any particular filter generation scheme so that comparison of the various options is possible. It would be highly desirable for the measure of quality to directly measure, in some way, the utility of the filter to its task of reversing the effects of the room acoustic on the signal to be reproduced.

Typically, comparisons of filter performance are made on the basis of mean squared error, the driving force of LMS adaptive systems. This solution has the advantage of being convenient to implement, since the value can be extracted direct from the adaption loop. Different methods, however, calculate the error figure used to drive the adaption slightly differently and thus using the error as a basis for comparison is questionable. (For instance, the error calculated during the application of the "filtered-x" algorithm is affected by the accuracy of the room model and thus is not simply related to the accuracy of the filter.) Even if a standard method of calculation were devised the problem remains that this measure has no direct, real-world, significance.

Adaptive filters are often used by the telephony industry to cancel either the effects of slight mis-matching on long-distance cables of audio/visual conferencing applications. The performance of the filters is typically measured in terms of their Echo Loss Return Enhancement (ELRE) [64], which is defined in terms of the function of the filter. (In teleconferencing applications the filter is attempting to prevent the output of the speakers at the receive end from being re-transmitted after it has been picked up by the microphones at the receive end. ELRE is defined as the ratio of the average power of pure echo at the send filter input to the average power of the residual error at the send filter output.) Some similar performance measurement is desirable for the HiFi application.

Figure 7.5: Echo Cancellation for Teleconferencing

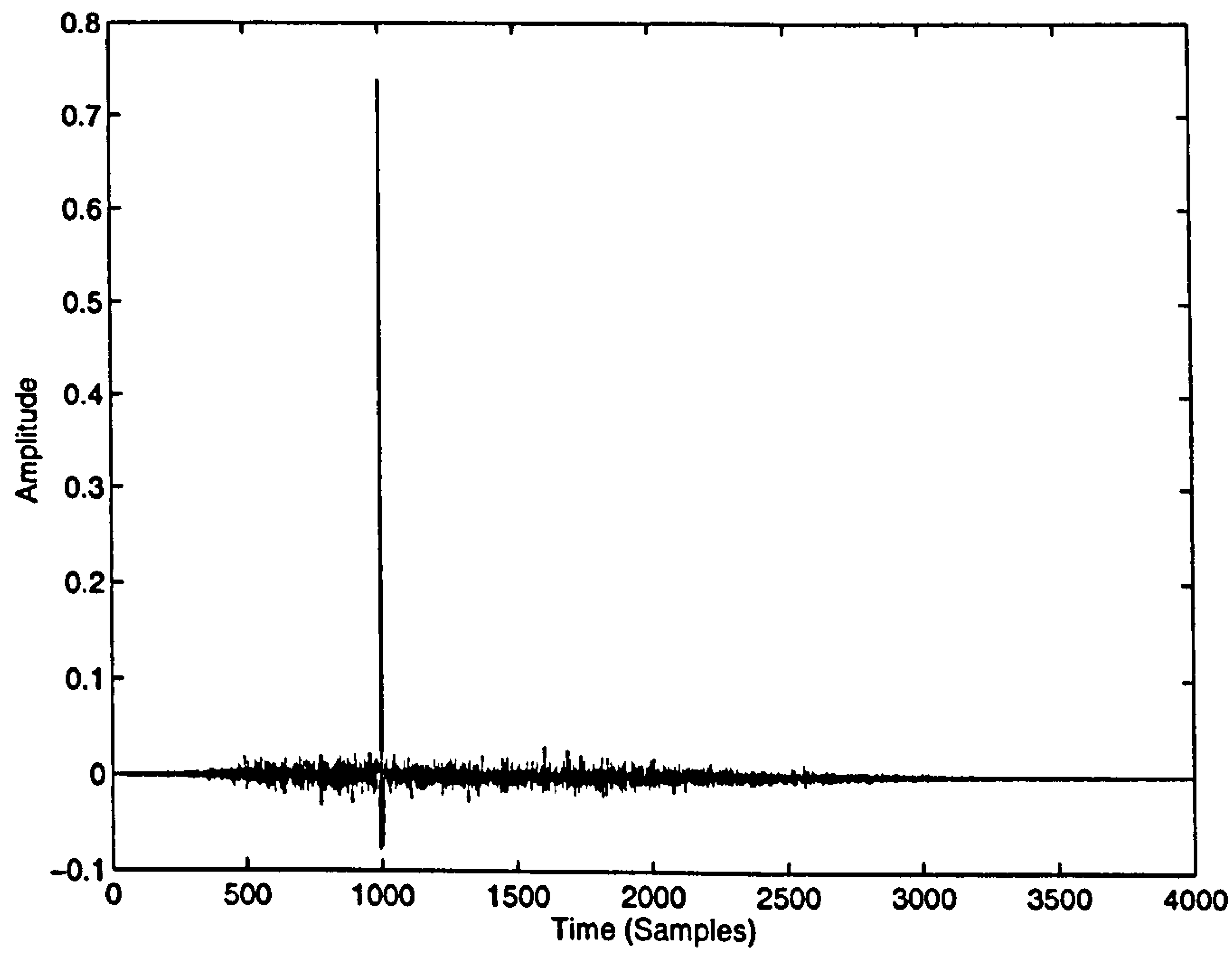


The problem is different for HiFi reproduction and thus it is not possible to make use of the ELRE measure. It is possible to measure the impulse response of the system including the proposed filter in all cases by using PRBS techniques. Using this and the fact that the desired impulse response is known, since that is the adaption target, then some general measure of effectiveness may be devised.

Examination of the measured system impulse response reveals a delayed impulse, but both before and after this desired response will be "noise". (Note that while these imperfections appear as noise on the impulse response they are consistent unless either the filter or the room impulse response is changed. Thus the effect of this "noise" on the output of the system is deterministic and signal related.) A typical response is reproduced in figure 7.6. Listening to a signal processed by an impulse response similar to that in figure 7.6 gives the impression of additional reverberation if the levels of the imperfections are significant, though it sounds odd due to the unnatural balance of decay, length and density. If the signal is such that there is a short silence or a significant change in level then the effects of the imperfections before the impulse are highlighted, the apparent "pre-reverberation" is unusual, and highly artificial sounding. Because it lacks natural analogue the levels of this pre-reverberation are particularly significant to the system performance. The pre-reverberation is due entirely to the "noise" on the system impulse response occurring before the desired impulse and thus measuring this noise will produce a measure of the levels of pre-reverberation that will occur when a signal is passed through the system.

In general the "pre-reverberation" may present a particular problem since music is generally composed of signals with short rise time and rather longer decay periods. It is difficult to make comprehensive tests since very significant amounts of time are required to create test signals with the available equipment.

Figure 7.6: A Typical Filtered System Impulse Response



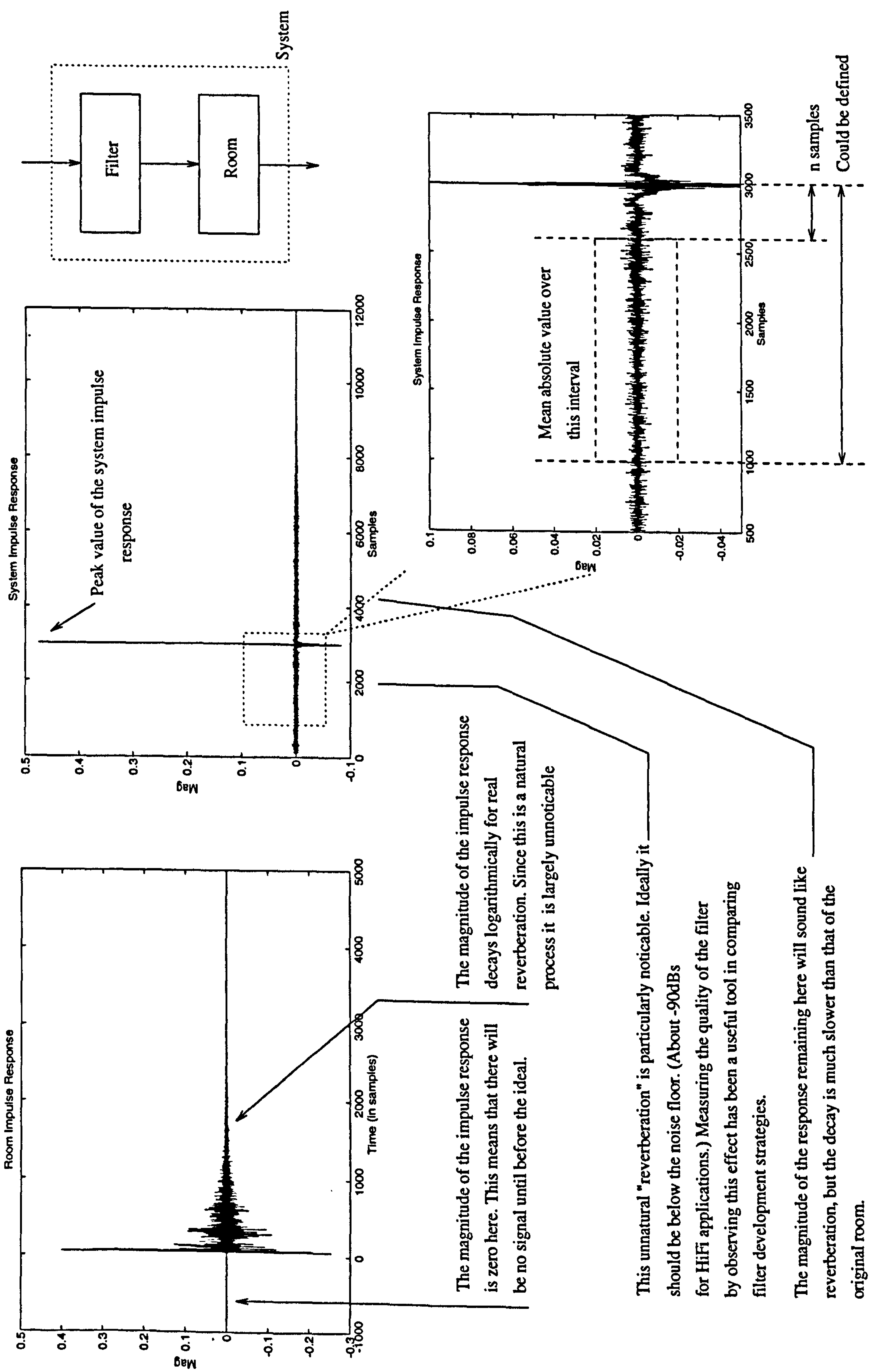
The quality of the filter could also be assessed in the frequency domain by Fourier-transforming the impulse response, but this is not strictly valid since music signals are not continuous. In practice the time-length of the impulse response is such that much of the impulse response "noise" will be heard as reverberation or possibly even echoes rather than tonal alteration. Thus some measurement of the levels of this reverberation is likely to be preferable to measuring improvements to the frequency response of the system. Note that the SigTech system uses a different rational, the designers have chosen only to attempt to deal with early reflections because these affect the tonality of the system response and ignore the later reverberation effects.

In general it would seem that if the "noise" in the system response was below the real noise floor then the signal due to an element of the system impulse response noise should be inaudible. Given that the noise floor of a typical high grade HiFi system is likely to be better than 90dB down and it is difficult to achieve an ELRE of much greater than 40dB for telephone echo cancelling this is likely to prove difficult.

The actual measure of filter quality initially adopted is to find the ratio between the peak and the mean of the "noise" each side of the desired impulse, see figure 7.7. The means of the "noise" either side of the desired peak are kept separate, to give the Pre-Peak Mean Ratio (PPMR) and the (post) Peak Mean Ratio (PMR), where the PPMR is the more significant statistic. (Simplistically speaking, provided that the PMR is lower for the system with the filter than the original impulse response then the filter may be judged a success from this point of view. As noted before the different characteristics of the post-peak "noise" when compared to natural reverberation mean that in practice lower levels are likely to be required.) In a normal stereo setup the PPMR of the system without correction will be zero and thus adding a filter will certainly degrade the original system performance by this measure.

The PMR and PPMR measures are convenient since they can be measured independently of architecture and sample rate, they can be measured for both real systems and simulated environments and values returned have some real-world significance. The PMR and PPMR for figure 7.6, which was calculated by a high accuracy method, are 46.7 and 47dB respectively. It is easy to see that this accuracy is comparable to the 40dB figure given for telephone ELRE, and is far below the 90dB figure required that has been suggested for HiFi applications.

Figure 7.7: Peak Mean Ratio Measurement



In practice the previous estimate of the impact of the impulse response noise could be conservative; consider that *each* element of the system impulse response that is inaccurate will contribute a continuous copy of the music signal delayed depending on its position in the system response then the real "noise" is the sum of these inaccuracies rather than the mean. Adjacent inaccuracies will not produce individually audible signals, over this short interval it is probably better to consider the music signal relatively continuous and thus the effect can be seen clearly if the inaccuracies are Fourier transformed.

Given that copies of a signal delayed over about 60-80ms will be heard as discrete signals rather than fused with the original [50] *pp 144, 146* [22] *pp325, 332, 345, 374* then any "noise" in the system impulse response over this length of time from the ideal impulse will tend to cause a discrete signal. Returning to the example system, figure 7.6 then any portion of the response more than 30 samples from the ideal will cause apparently discrete signals. (30 samples is approximately 70ms since this system was based on a response initially measured at 44.1kHz, but subsequently decimated by a factor of 100)

Since any copy of a signal separated in time by more than about 70ms will be heard discretely from the original then a useful estimate of the impact of the reverberation could perhaps be obtained by comparing the frequency response of the entire impulse response to that of a section starting more than 70ms after the main impulse.

If a section of the system impulse response at greater than 300 samples after the ideal impulse is taken and the resulting mini-impulse response is Fourier transforming a frequency domain representation of the pass-band for the filter that these samples comprise is produced, see figure 7.8. (In practice no discrete signal is heard since the pseudo reverberation that the impulse response "noise" causes is continuous.) Effectively there will be a copy of the input signal filtered by this filter that is delayed by over 700ms, ten times the value that causes the result to cease to fuse.

By comparing the pass levels of these two "filters" an estimate of the impact of the impulse response "noise" can be obtained. Figure 7.8 is the frequency response of the section of the impulse response occurring long after the main impulse and figure 7.9 is the Fourier transform of the entire system impulse response. When compared to the Fourier transform of the entire system (figure 7.9) it can be seen that this "second copy of the signal" is scarcely 20dB down on the the desired signal. The difficulty of making useful comparisons when attempting to account for these complexities means that on the whole the PMR and

PPMR will be used for filter performance measurement. The point of this analysis is to emphasis that even if a filter were developed that achieved a PMR and PPMR of 90dB the performance may still be less than satisfactory in practice. Even if higher standards of PMR and PPMR are required these parameters are expected to be useful measures of performance,

Figure 7.8: Frequency Response for End Samples

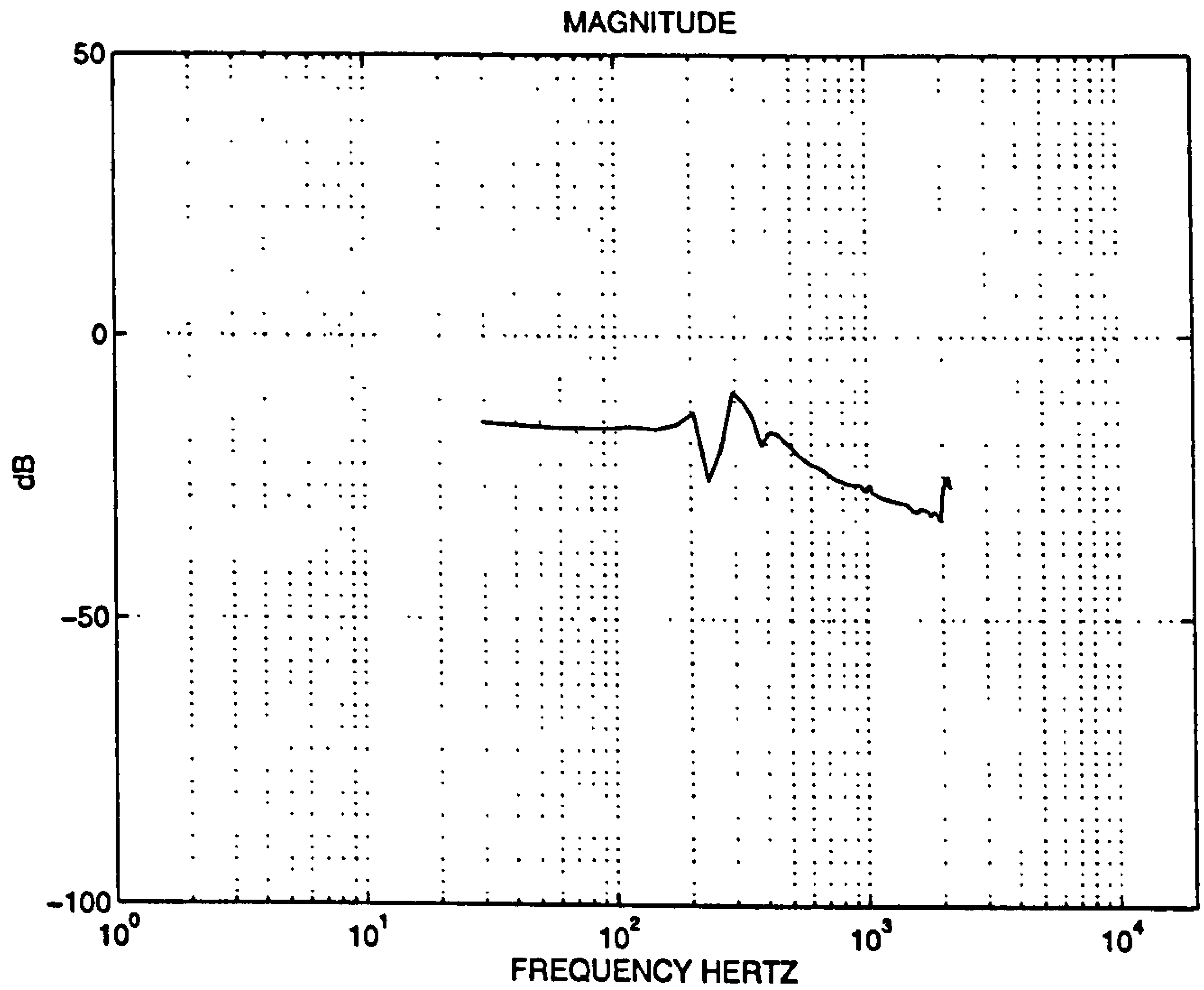
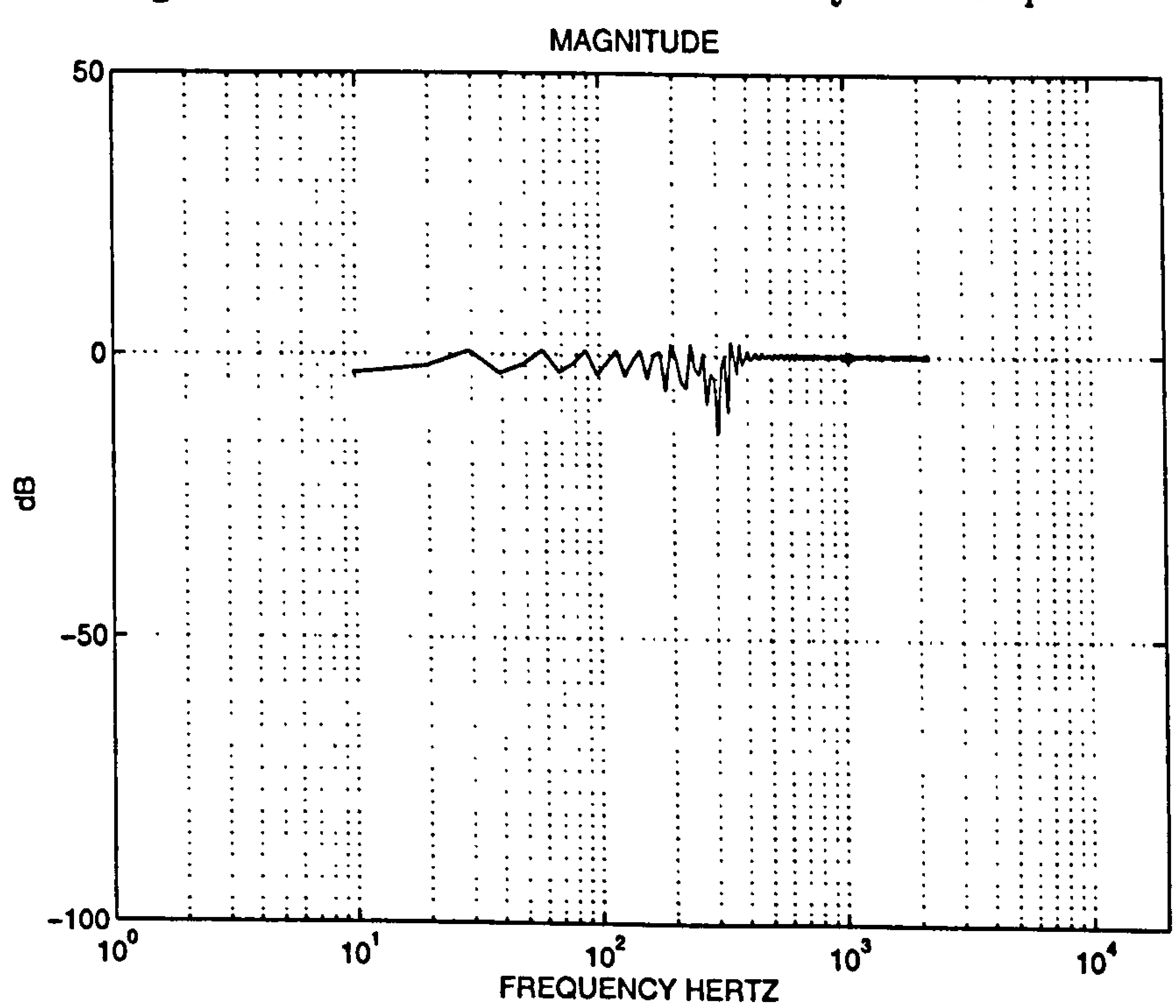


Figure 7.9: Fourier Transform of the System Response



7.5 Architectures for Adaptive Inverse Filtering

In this section the differences between Widrow's "filtered x" structure [62] pp288-297 which was favoured by Nelson *et al* [44], [46], [43] and an alternative "pre-filtered" structure are discussed; the latter offers considerable improvements in performance when applied to room acoustic compensation for HiFi sound reproduction systems. The pre-filtered structure was briefly considered by Widrow [62] pp281 but has not been applied to this problem. It is possible to extend this concept to stereo reproduction and cross-talk cancellation by a similar method to that proposed by Nelson *et al.* [44] for the "filtered x" algorithm. The proposed alternative is compared with the "filtered x" algorithm in terms of the adaption rate, computational complexity, stability and convenience.

7.5.1 Widrow's "filtered x" algorithm

Figure 7.10: The "Filtered x" algorithm

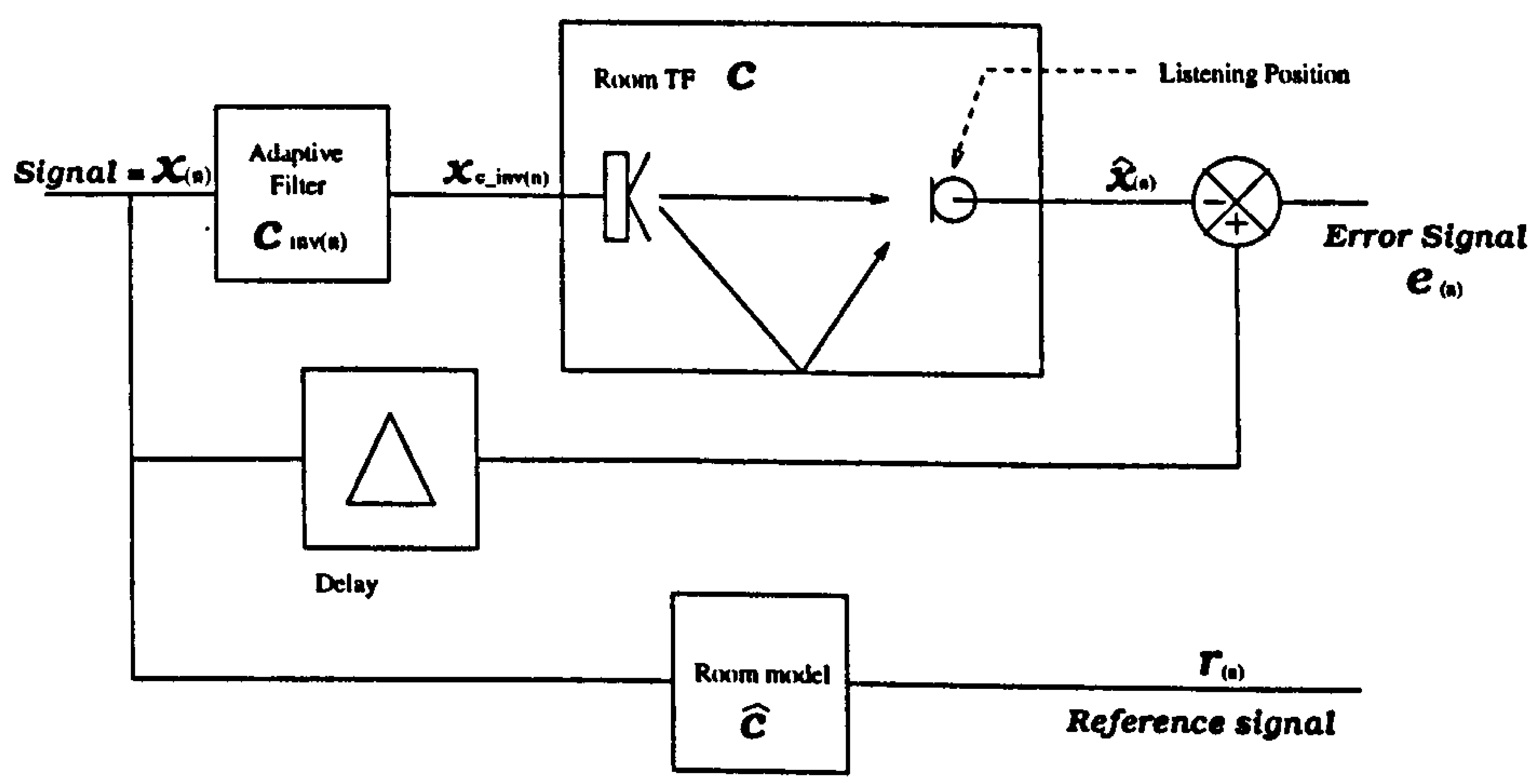


Figure 7.10 gives a typical application structure for an adaptive room acoustic equalization system after Widrow. This style of solution has been extended to multiple channels by a number of different people to cope with stereo signals, improve performance etc.

The operation of this algorithm as follows; the signal first passes through the adaptive filter and then the room where it is picked up by the sensor. Next the value sensed in the room is compared with a delayed version of the original signal, which is what would have been sensed if the filter had been ideal. The difference is used to alter the filter weights and thus improve the filter. The weight update procedure is slightly different from the NLMS algorithm; since the order of the room and the filter have been reversed it is necessary to use the reference signal $r_{(n)}$ instead of the filter input. The reference signal is generated by making use of a model of the room, $\hat{c}$ that needs to be measured initially.

The three principal problems with the "filtered x" algorithm are:

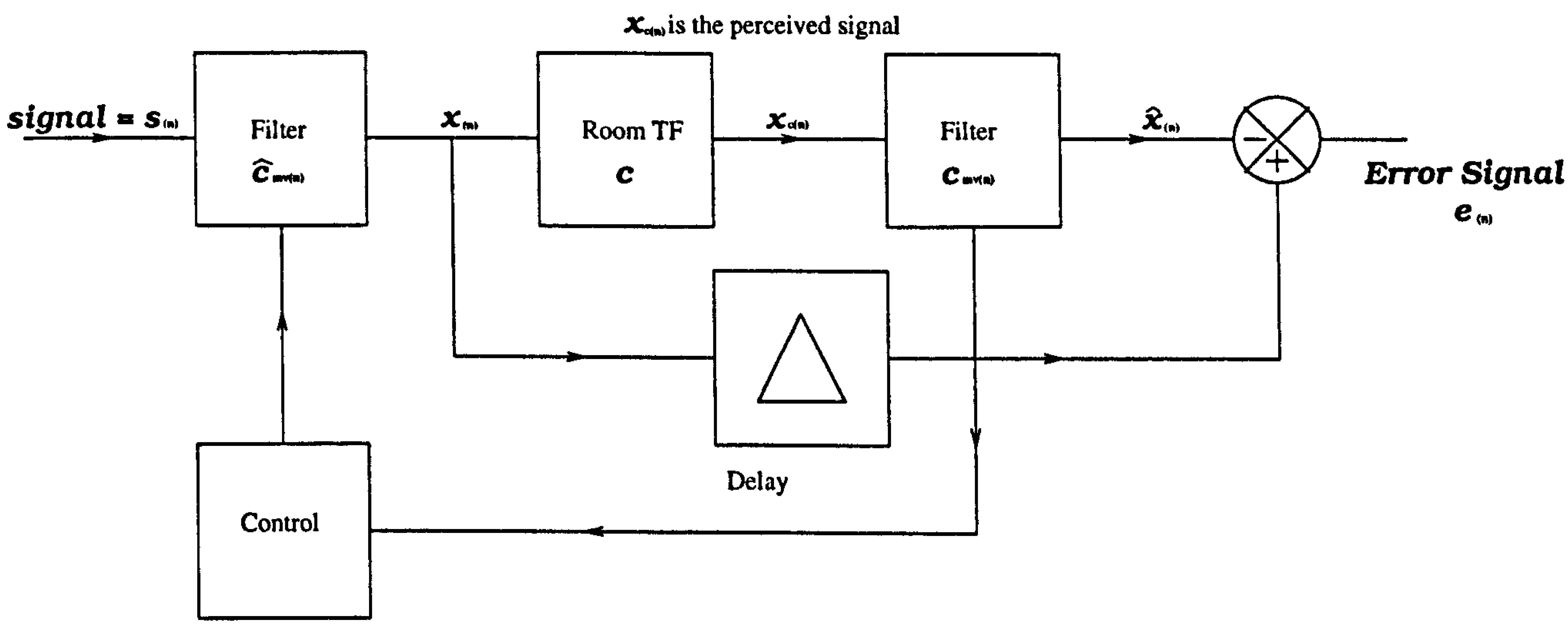
- Transfer function model required. An accurate model of the room transfer function is required, this in fact poses a two-fold problem.
  - Transfer function measurement. The transfer function can be measured by PRBS methods but the noise pollution produced as a consequence would probably be intolerable in a HiFi environment. (PRBS signals are very wide band and random signals, which probably contributes to their subjectively unpleasant sound. It seems unlikely that the type of people who would make significant investments in their audio reproduction apparatus would accept such a signal being reproduced over their equipment and tolerate the pervasive noise produced.) It is possible to extend the architecture to include an adaptive estimator of the room transfer function but this requires additional processing power and will further adversely affect the adaption time of the system especially when the accuracy required of the model is considered.
  - Accuracy of the transfer function. If the transfer function model is not accurate the adaption speed will be adversely affected. As the phase accuracy decreases the adaption constant (β) has to be decreased to maintain stability; at ± 90 degrees error the system becomes unstable [3].
- Distortion during the initial adaption. When the system is started there is likely to be considerable distortion of the signal, until a reasonable approximation of the inverse filter is arrived at. It would be possible to use a wide-band training signal to reduce this time but this would again be intolerable from the point of view of noise pollution.
- Adaption time. Preliminary tests have shown that obtaining a sufficiently good solution for HiFi applications in a reasonable amount of time is likely to be a problem.

7.5.2 The Pre-filtered Algorithm

This solution is based around a well understood adaptive process. It boasts an improved adaption rate, need not cause distortion of the signal during initial adaption and does not require any knowledge of the room transfer function but a substantially more complicated control system is required. The additional processing power required is negligible.

Initially the filter $\hat{c}_{inv(n)}$ is set to an impulse so that all of the signal is passed. Then, as time progresses the filter $c_{inv(n)}$ learns to be a good inverse of the room. In this solution the room and filter are in the normal order and thus there is no need for a room model for the adaption process. All that is required is the error, which is calculated in the normal way and the signal $x_{(n)}$. When the filter is judged to be sufficiently mature a copy of $c_{inv(n)}$ is transferred to $\hat{c}_{inv(n)}$ and then correction begins.

Figure 7.11: The Pre-filtered Algorithm



7.5.3 Performance Comparison

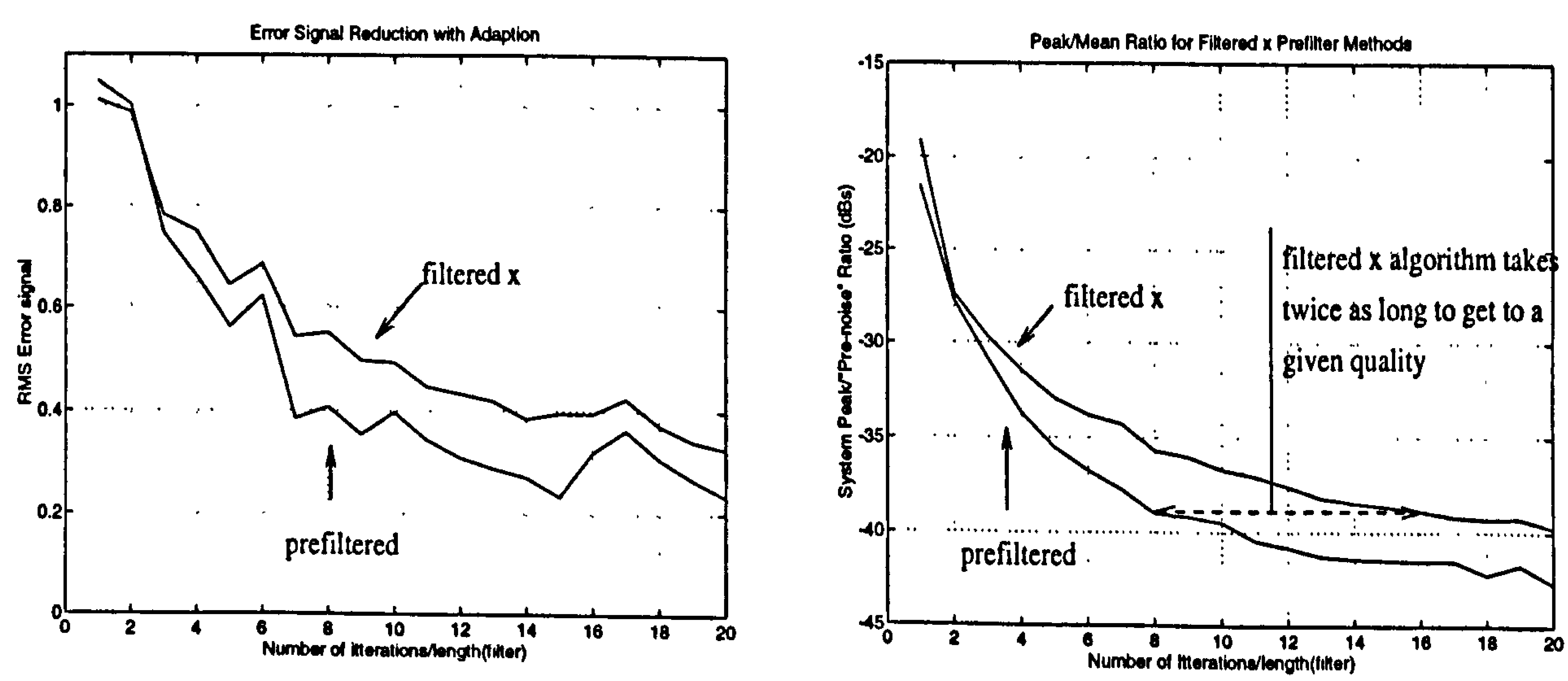
Adaption Rates

It is difficult to quantify adaption rates since the differences in structure means that the systems need never arrive at exactly the same answer. Some indication of the relative "quality" of the filter is required. The following measures can be used:

- Error signal. A figure derived from the error signal is often used to indicate the progress of a filter. In this case it may not be a good indication since the architectures differ markedly. Figure 7.12 includes a plot of the square root of the mean square of the error signal values over each successive period of L iterations, where L is the length of the filter. It can be seen from this plot that the pre-filtered algorithm is superior, but it is difficult to make a judgment of the relative quality. Note that both processes were running with their optimum adaption constants and since these results were simulated the filtered x algorithm was able to have a perfect model of the room.

- Peak to mean "noise" ratio. If the the system impulse response is calculated, by convolving the room impulse response with the filter developed to compensate for it, the ideal case result would have only one non-zero sample. Practical results resemble this situation but the remaining locations are contaminated with "noise". When convolved with the music signal it is possible to hear this "noise" on the impulse response as additional "reverberation" of signal, both leading or lagging that desired. This undesirable trait can be quantified if the ratio of the peak to the mean noise over a defined interval of the system impulse response is measured. Since music is characterized by fast attack and slower decay any "noise" on the system impulse response before the ideal impulse will be particularly audible hence this alone is used in the measurements. Figure 7.12 includes a graph showing how the value of this noise reduces as the adaption process proceeds. From this graph it can be concluded that the pre-filtered algorithm adapts approximately twice as fast. In practice the advantage will be greater than this as explained when the stability limitations are considered.

Figure 7.12: Performance Comparison



Computational Complexity

The additional calculations required for the prefilter of the pre-filtered algorithm are completely offset by the lack of a room model in this solution thus the only increase in processing power is that required by the control process. The function of the control process is to transfer a copy of the filter developed to the prefilter when it is considered mature and in a way that makes the transfer inaudible; this is unlikely to be a significant computational requirement when compared to the main process. (A slow morphing process is probably the best way to change between filters. This would not work well if the system transit time had changed, for instance the listening point had moved closer to the speakers. The desired response could be altered, by changing the modelling delay, to hide the effect of changes in the transit time from the listener.)

Stability

The core of the pre-filtered algorithm is a well understood normalized adaptive process that is known to be stable with the sole provision that the adaption constant β is below about 1.2; 0.6 gives good results.

In contrast to this the filtered x algorithm, as already noted, has stability and the value of the adaption constant governed by the quality of the room model. It will never be possible to run the algorithm at the maximum adaption rate because if the room changed sufficiently the system would become unstable. (For instance, if the listening point were moved, very significant changes in the room impulse response could be caused.) At the extreme HF of a full bandwidth signal the wavelength is a mere 17mm, which gives a 90 degree phase change in 4.3mm. Realistically it will not be possible to guarantee that the microphone does not move by this kind of distance. In [3] the optimum value of β is calculated to be in the range $0.15 \leq \beta \leq 0.2$ but in [2] the practical range for stable convergence was found to be $0.01 \leq \beta \leq 0.05$ for a much reduced bandwidth.

Convenience

The convenience of having the adaption loop outside of the listening part of the process should not be under-emphasised. It allows the filter adapt without subjecting the listener to the early results which would be expected to be poor. The prefilter algorithm also allows the filter in use to be "toned down" which would not be possible if this was in the adaption loop. SigTech found it necessary to "tone down" their fixed filters to make the results more acceptable.

Although the filtered x algorithm is quite tolerant of background noise the pre-filtered algorithm guarantees that the filter in place will not be corrupted by short term noise signals such as closing doors.

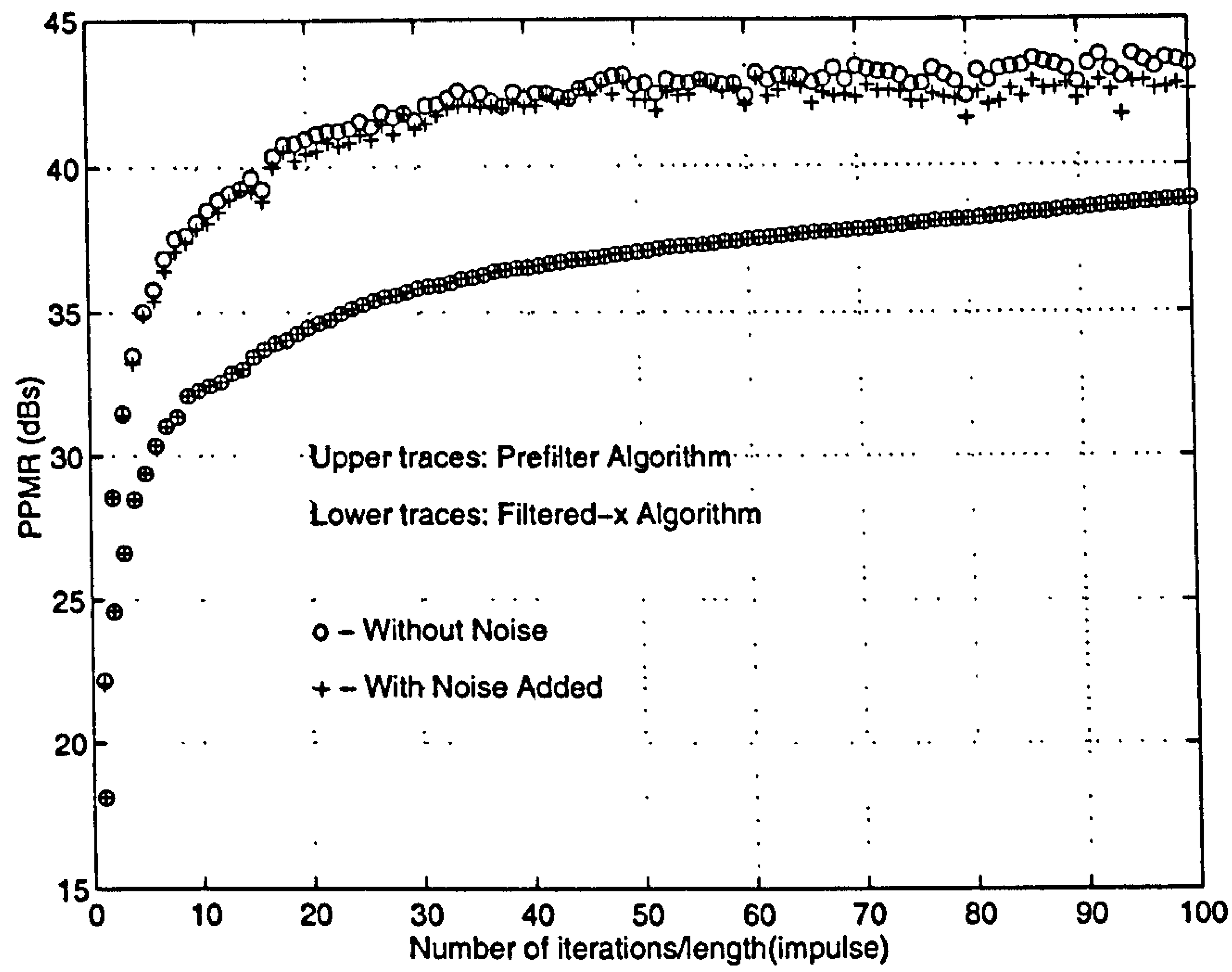
7.5.4 Noise Performance

The chief advantage given by Widrow for the filtered-x algorithm is much improved performance in a noisy environment. The filtered-x algorithm has the advantage since any noise in the real system is present directly at the error calculation stage rather than being convolved with the inverse filter response. The effects are therefore less damaging.

As a test noise was added to the signal coming out of the simulated room for each of the architectures at -30dB roughly relative to the signal level. (In practice the noise level was somewhat higher that this due to an implementation technicality.) Values of beta were selected for each system that were half of the reported practical maximum for a stable system. (For the filtered-x system a value of $\beta = 0.03$ was used and for the pre-filtered system $\beta = 0.6$. The resulting PPMR/adaption plot is given in figure 7.13. It can be seen that though the pre-filtered method suffered far more from the noise it still offered a much better result even after long adaption. (Note that the difference between the PPMR for the filtered-x algorithm with and without noise was very small.)

The value of noise applied was deliberately selected at a very high level compared to the values expected in a real system so that there could be no chance the effects of the added noise could be lost in the noisier test system used here. (The test system PMR is not nearly good enough to have practical application in HiFi due to the very short filter and very low number of taps.)

Figure 7.13: PMR Development as Adaption Progresses



7.5.5 Conclusion

For a real system it would be impractical to make use of the simplest form of Widrow's filtered-x algorithm, as was suggested by Nelson *et al*, since this requires a room model that would have to be measured in advance using PRBS techniques with the attendant and intolerable noise pollution. (It is possible to measure the room impulse response by other methods but the performance is likely to be inadequate.) If this system were chosen then many of the advantages of an adaptive system would be lost since movement of the measuring microphone would result in the system going unstable.

The only practical advantage that might be gained by applying Widrow's filtered-x algorithm is that its performance in the face of external noise is better, but even this did not offset the performance advantage of the pre-filtered algorithm.

Widrow suggested more complicated forms of the filtered-x algorithm [62], but these would require additional processing. It is true that the room impulse response does not have to be measured for these more advanced algorithms and thus some of the important disadvantages are addressed but the price for this added flexibility is likely to be even slower adaption.

The pre-filtered algorithm has none of these disadvantages, it adapts quickly, requires no room model and will remain stable even if the room impulse response changes dramatically. It is even possible to include control that will apply the inverse filter sensibly, ie. not until it is mature and only while it remains valid, that is the error rates remain low.

7.6 Possible Improvements to the Adaptive Filter

It has been established that the "pre-filter" architecture is superior to the "filtered x" alternative for room acoustic inverse filtering applications, however, so far only the basic NLMS algorithm has been assumed. Achieving sufficient performance even when using the "pre-filter" architecture is expected to be a problem. There are a large number of related alternatives to the NLMS algorithm that offer different balances between the processing power requirement, convergence rate and final accuracy which warrant consideration. Initially the more minor changes will be considered in detail but more radical options will be discussed later.

There are a number of slight modifications to the base NLMS algorithm that have been suggested in the context of the telephone echo cancelling problem. (The telephone problem has received greater attention since it is of greater commercial interest.) The telephone echo cancelling problem is in many ways similar to room acoustic compensation, but there are a number of differences that mean transferring the technology is not completely straight forward. The options that will be considered under this heading are:

- Step Weighting
- Time Weighting
- Oversampling
- Multiple Channels
- Over-adapting

In addition to these options there are alternative methods that offer a completely different balance to the adaption problem. In practice there are few, if any, methods that can compete with the power of the LMS algorithm in pure form and thus the more interesting of the methods considered here are hybrids. The Methods that will be considered will be:

- Hybrid Newton/NLMS method
- Frequency Domain Methods
- Fast Newton
- Wavelet Based Methods

7.6.1 Modifications to the Pre-filter Algorithm

Step Weighting

Makino *et al* published a paper describing a method for improving the convergence rate of echo cancellers designed for telephony application [36]. (Makino also published a very similar paper in the IEEE Transactions [37].) In these papers they claimed to be able to double the convergence speed of the NLMS algorithm for a modest additional processing requirement by considering the statistics of the room impulse response.

Briefly, it is important to consider the function of an echo canceller designed for telephony use. The system is a loudspeaker and microphone in a room whose purpose is to give a hands-free fully duplex conferencing facility to the user. Thus it is important that the loudspeaker and microphone are operational all of the time, which inherently leads to the possibility to feedback or howl-round. (In less extreme cases there would still be the disconcerting effect of the echoes and the loss of intelligibility.) The echo canceller is designed to remove as much of the signal produced by the loudspeaker from that picked up by the microphone as possible. In order to achieve this aim an adaptive system is employed to identify the room transfer function.

Since the aim is to produce a model of the room transfer function in real time the convergence speed of the algorithm is critical. Makino *et al* proposed that by considering the statistics of general room impulse responses it should be possible to add some general information and thus improve the adaption speed. It appears that the mean amplitude of a room impulse response typically reduces exponentially with time and thus it would be reasonable to expect the FIR model's weights to follow suit. (The weights of an FIR filter are identical to the samples of its impulse response.) Thus the tap update equation becomes:

$$\mathrm{h}_{k+1} = \mathrm{h}_k + \mathrm{A}\frac{e_k}{||\mathrm{x}_k||^2}\mathrm{x}_k \tag{7.2}$$

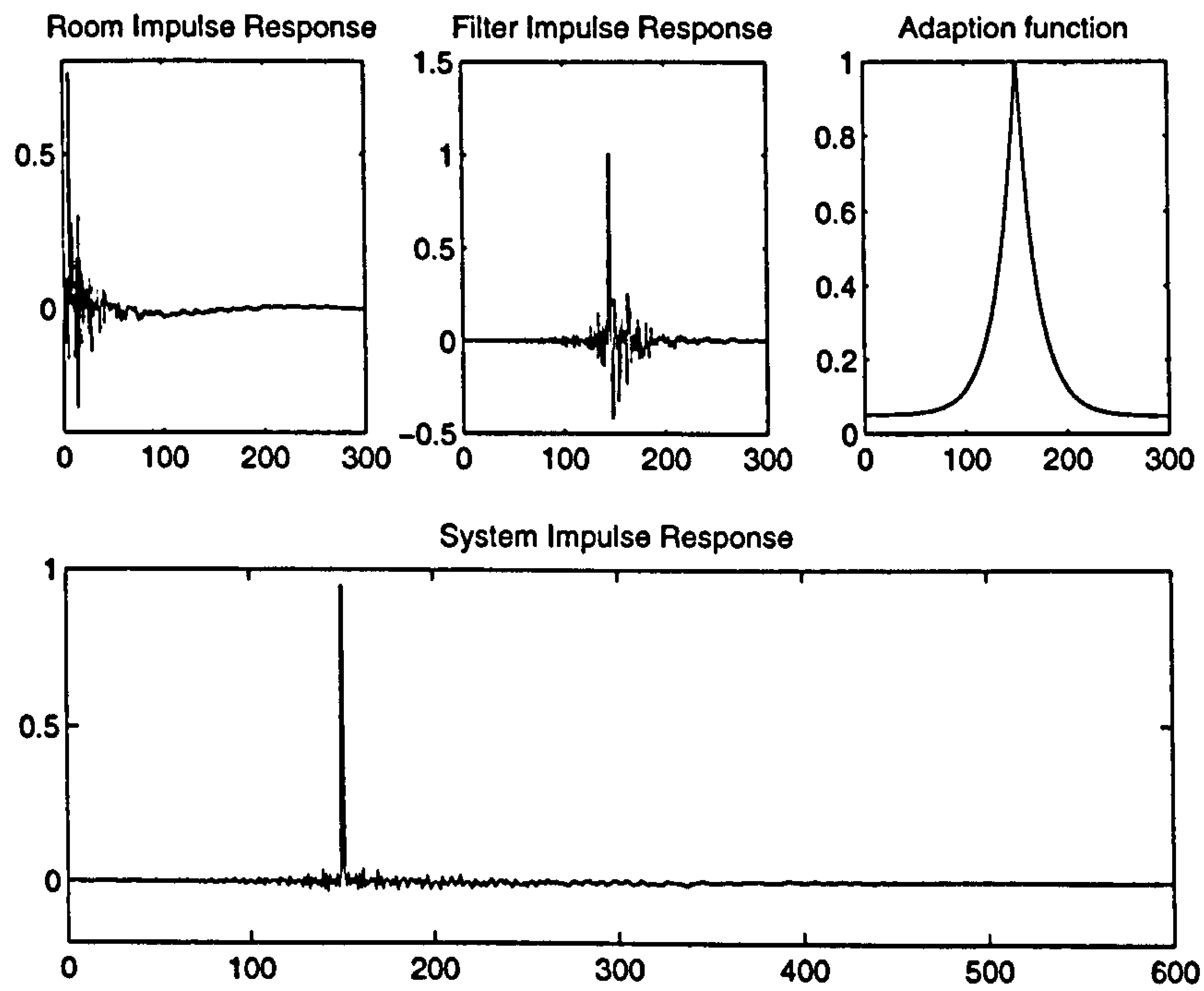
Where the weight matrix is of the form:

$$\mathrm{A} = \begin{bmatrix} \alpha_1 & & & 0 \\ & \alpha_2 & & \\ & & \ddots & \\ & & & \alpha_n \end{bmatrix} \tag{7.3}$$

Makino *et al* further suggest that the step weighting should change by powers of two, a very effective measure of reducing the computational load required for this enhancement. At the end of the paper practical results are given that indicate a substantially faster convergence for the step-weighted system, though it is not clear exactly how the error was calculated in each case.

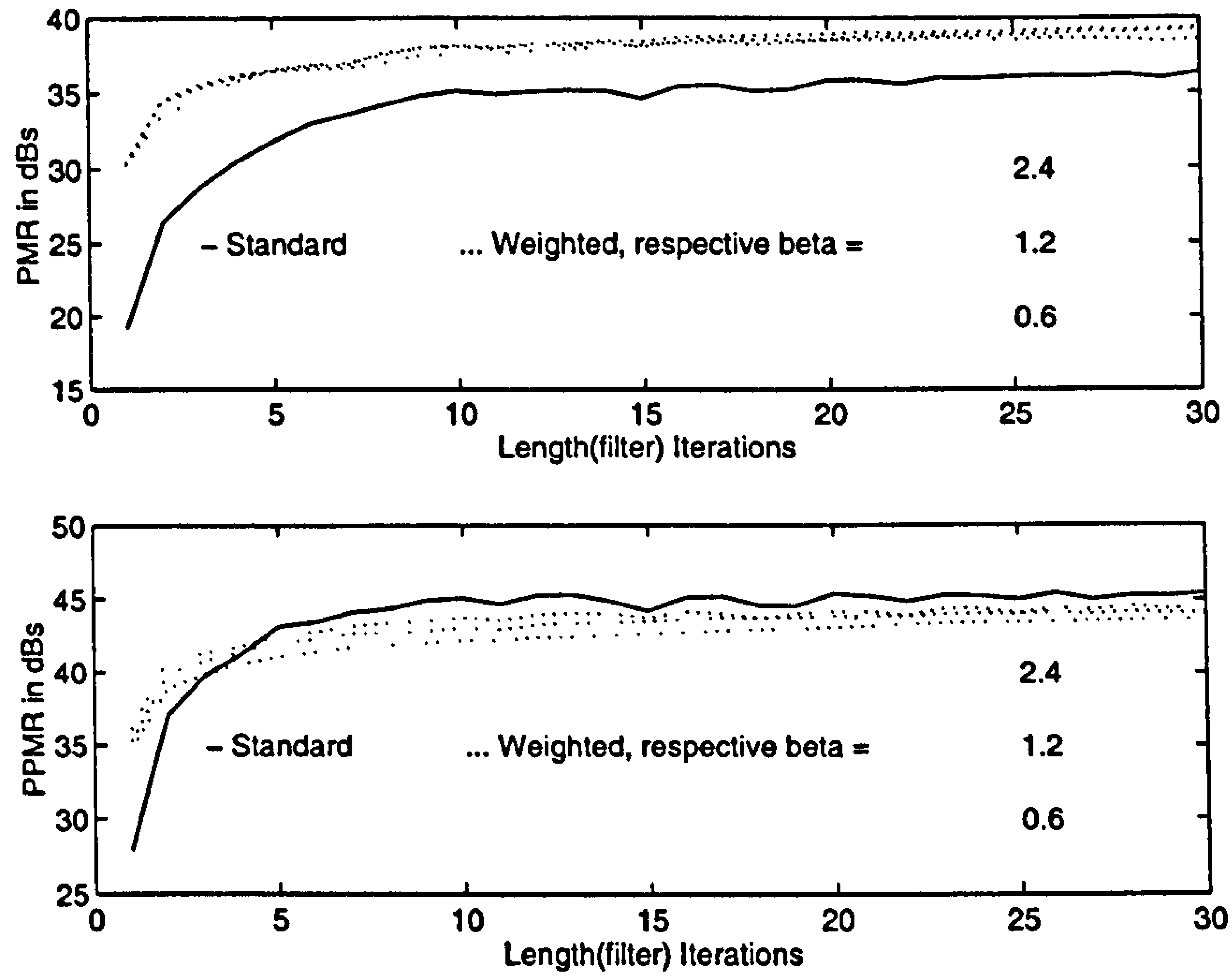
The room acoustic correction problem is a little different, but the application of a similar technique should be possible provided that some statistical representation of the expected filter could be produced. In practice the variance of the room acoustic filter indeed shows a predictable pattern and thus it is possible to apply this technique. The filter's variance approximately follows an exponential decay either side of the delay thus a window was constructed using parameters that gave a good match to the target system. A sample result is given in figure 7.14

Figure 7.14: Filter Produced Using Variable Step Weighting



The new algorithm was compared to the pre-filter algorithm for a variety of values of β. (Note that only $\beta = 0.6$ was used for the NLMS algorithm since this is known to represent a good compromise.) It is clear from the results (figure 7.15) that the new algorithm consistently offers a better convergence rate initially and may produce better results generally. There is little doubt that initially applying this algorithm will improve the convergence rate but it may be better to change back to the NLMS algorithm after, say, 10 length(filter) iterations to ensure repeatable final convergence.

Figure 7.15: PMR and PPMR for Original and "Improved" Algorithms



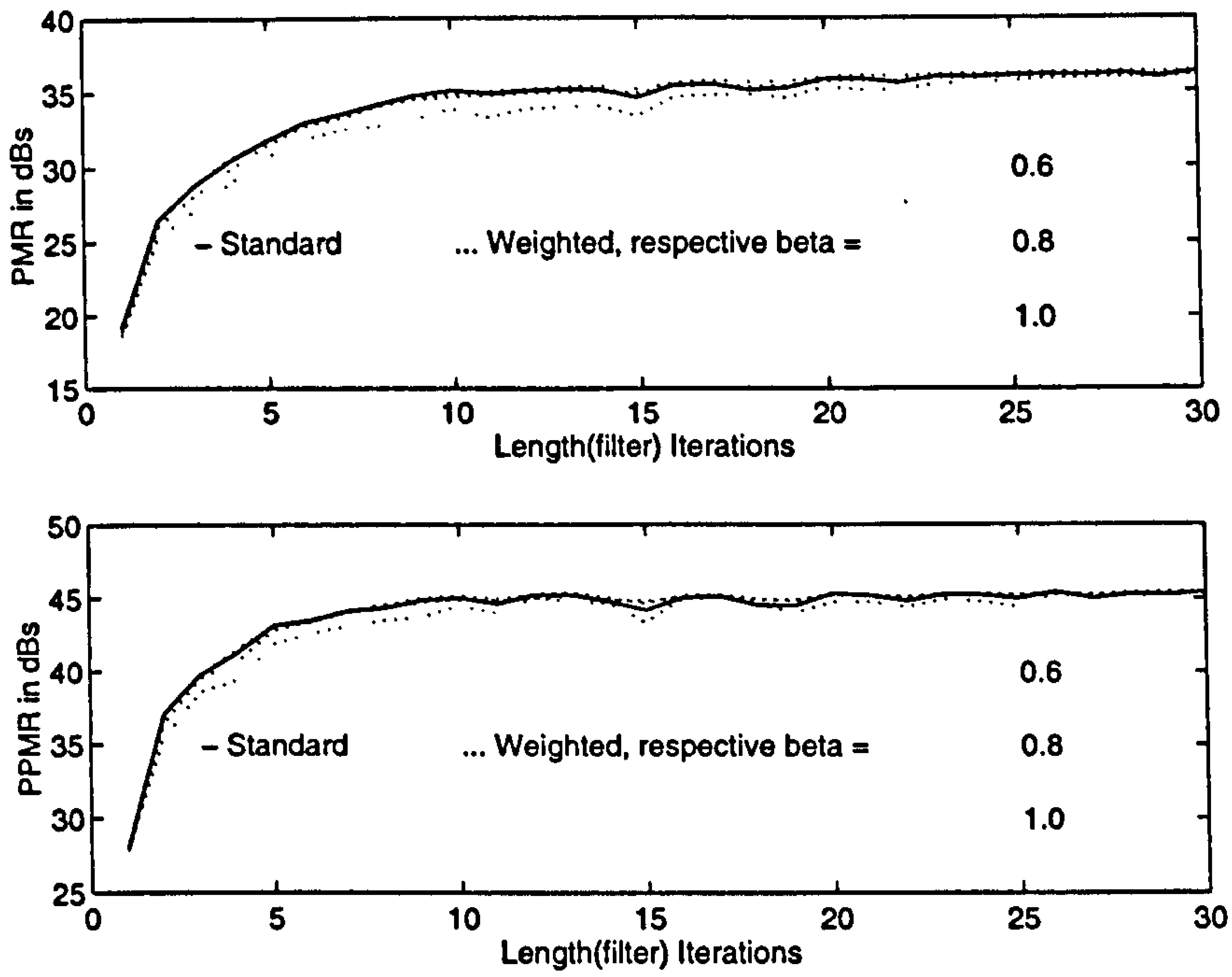
Time Weighting

The concept is to reduce the effective adaption constant β as the development of the filter progresses so that good use can be made of both the faster initial adaption of high β values and the lower mis-adjustment of lower ones. Time weighting is perhaps not a very good way of describing this technique, in practice it is far easier to adjust the error weight by measuring the approximate mean error. This process was suggested by Kwong *et al* in their paper [33], though again the focus of the paper was the telephony application.

The test process takes the mean-square of the error over each length(filter) iterations and uses this to update the error weight (effectively reduce β). It was found that a conditional statement was required to ensure that the weighting never exceeded 1.0. The inclusion of this additional limit ensures stability regardless of the value of the error.

Tests indicate little or no practical advantage to applying this method. Figure 7.16 shows that the PMR and PPMR do not reduce to significantly lower levels or do they reduce more quickly initially.

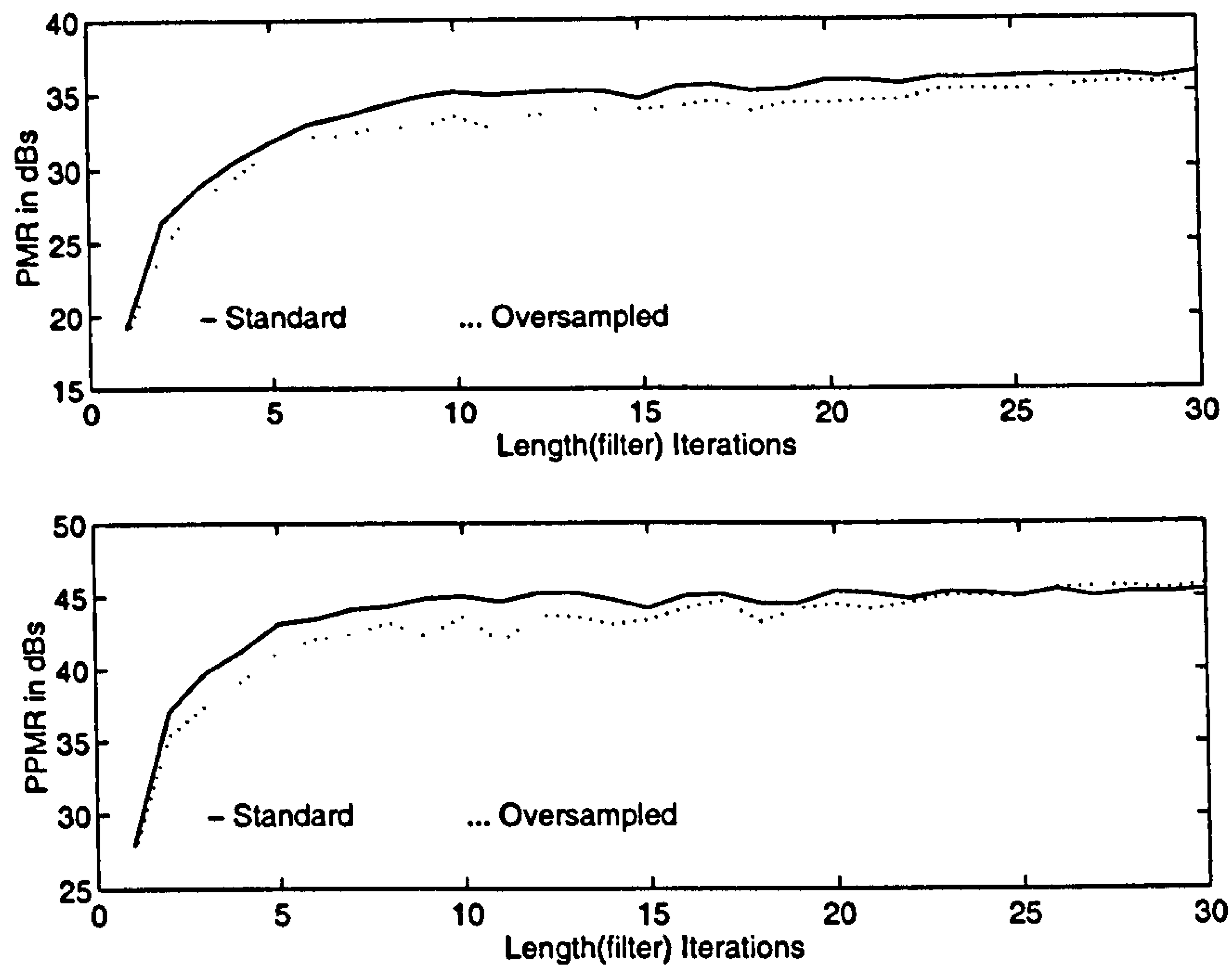
Figure 7.16: PMR and PPMR Comparison for Step Weighting by Time



Oversampling

It is possible to operate the inverse filter system at twice the bandwidth of the signal it is to process. The aim of this manoeuvre is to hopefully improve the accuracy of the filter produced by giving increased resolution from increasing the overall number of coefficients. In practice there is no apparent advantage in this procedure. Figure 7.17 shows the PMR and PPMR for two filters, one standard and the other oversampled by a factor of two. Limit calculations also do not indicate any potential gains from applying this process. (The process of making limit calculations is discussed in the introduction to adaptive processes and DSP.)

Figure 7.17: PMR and PPMR for Standard and Oversampled Filters



Multiple Channels

Room systems are very sensitive to small changes and thus a system with two loudspeakers, even if placed with apparent bi-lateral symmetry, will probably have quite different transfer functions from each speaker to the listening point. Allowing the NLMS algorithm access to both channels separately will therefore probably give much better results since the algorithm will chose the easiest route for each component of the signal to be corrected. This system cannot be applied where the listener would be able to detect the difference in the sounds produced by the two loudspeakers. This is not a severe limitation, referring back to the introductory chapters, Section 3, the stereo system will not produce directional information useful to the human hearing system below frequencies of about 1800Hz.

The use of multiple channels to improve the equalization was introduced to room inverse filter production at roughly the same time by both Nelson *et al* [44] etc. and Miyoshi *et al* [39]. The application of this idea requires more significant modifications to the NLMS algorithm. Considering the matrix method, see Appendix 9.1. If the single room transfer function (g_k)were to be replaced by two separate ones (g_{1k}, g_{2k}), one for each channel of the stereo system, and likewise two filters were used (h_{1k}, h_{2k}) then the convolution expression would now become:

$$\mathbf{d_k} = \begin{bmatrix} g_{10} & & & g_{20} & & & \\ g_{11} & g_{10} & & 0 g_{21} & g_{20} & & 0 \\ \vdots & g_{11} & \ddots & \vdots & g_{21} & \ddots & \\ g_{1L_g} & \vdots & & g_{10} g_{2L_g} & \vdots & & g_{20} \\ & g_{1L_g} & & g_{11} & g_{2L_g} & & g_{21} \\ 0 & & \ddots & \vdots 0 & & \ddots & \vdots \\ & & & g_{1L_g} & & & g_{2L_g} \end{bmatrix} \begin{bmatrix} h_{10} \\ h_{11} \\ \vdots \\ h_{1L_h} \\ h_{20} \\ h_{21} \\ \vdots \\ h_{2L_h} \end{bmatrix} \quad (7.4)$$

Where g_{27} is the seventh element of the second room transfer function.

By careful selection of the impulse and filter lengths it is possible to make this extended definition of the matrix $\mathbf{G_k}$ square, and thus amenable to direct inversion. Obviously the results of such an inversion are accurate inverse filters, accuracy in fact limited only by the precision of the arithmetic. It is expected that it is this situation that led Nelson *et al* to comment that there could be a filter length that would permit an accurate inverse [46]. (Miyoshi *et al* also commented that with careful choice of vector lengths an exact inversion of the Gk matrix was possible, but made no further comment.)

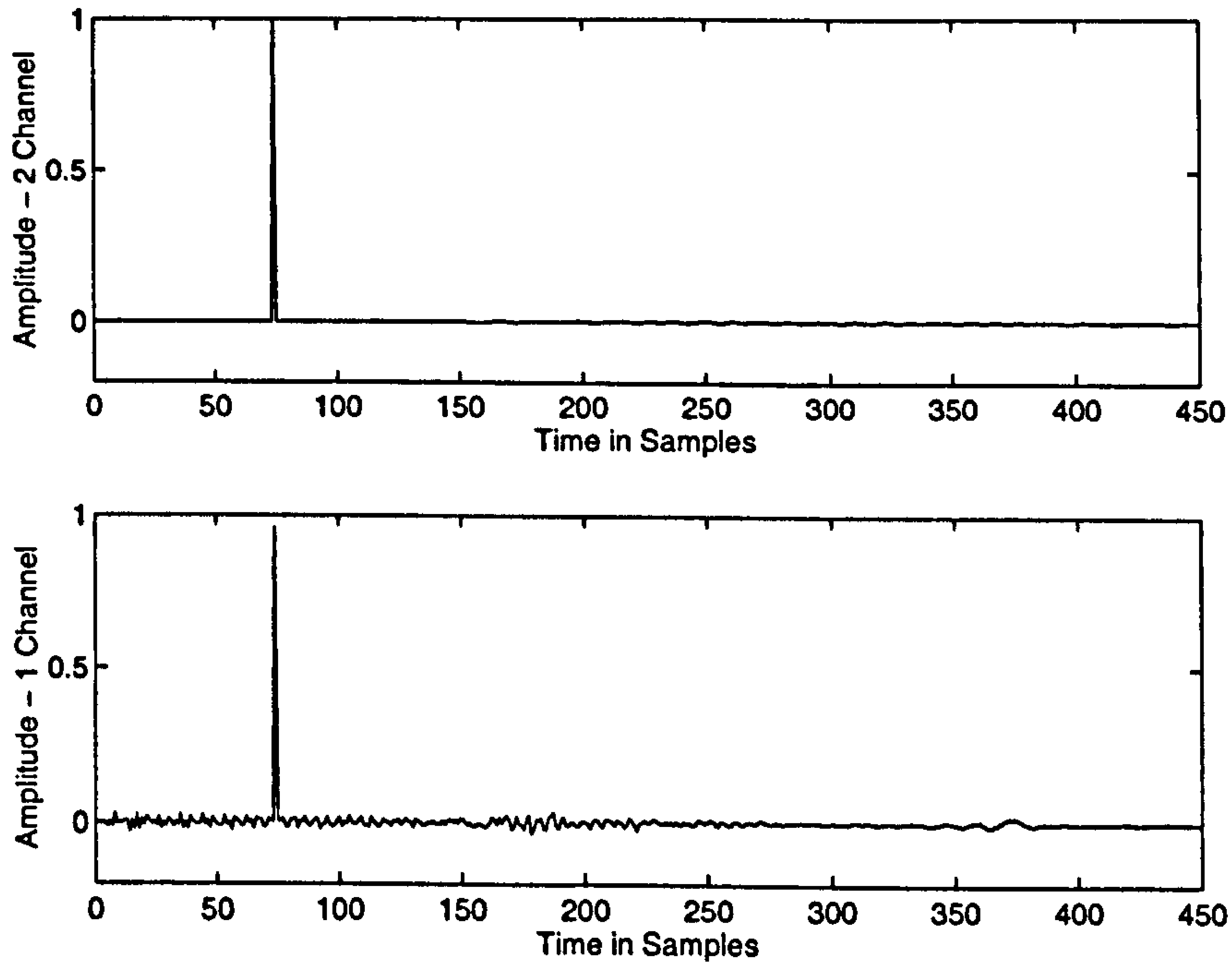
Unfortunately, in the real world, it is not possible to arbitrarily select the length of the room impulse response, it is an infinite series, and thus the exact filters that were calculated with $\mathbf{G_k}$ square are not exact solutions to the problem. The filters calculated are exact only when used with the truncated versions of the impulse response that was used to generate them.

The use of two paths may still prove beneficial, indeed it is to be expected that it would. If the impulse responses used for the tests are substantially longer than the proposed filters, say twice as long, then a more realistic expression of the problem may be obtained. It is not possible now to produce the inverse filters by simply inverting the $\mathbf{G_k}$ matrix, since this is not square and thus no inverse is defined, but a similar solution to that applied to the single path can be used to produce the optimum approximate solution. (It can be seen from examining the original and the enhanced convolution expressions that overall they are of the same form, though the definitions of the elements differ.) Using;

$$\mathbf{h_k} = \left[\mathbf{G_k^T G_k}\right]^{-1} \mathbf{G_k^T d_k} \quad (7.5)$$

will therefore produce the inverse filters for the two path system as readily as for the single path example. Using this technique inverse filters have been produced for a room using both the single channel and multiple channel techniques. The resulting system impulse responses are given in 7.18.

Figure 7.18: System Impulse Response for Single and Multi-Channel Filters

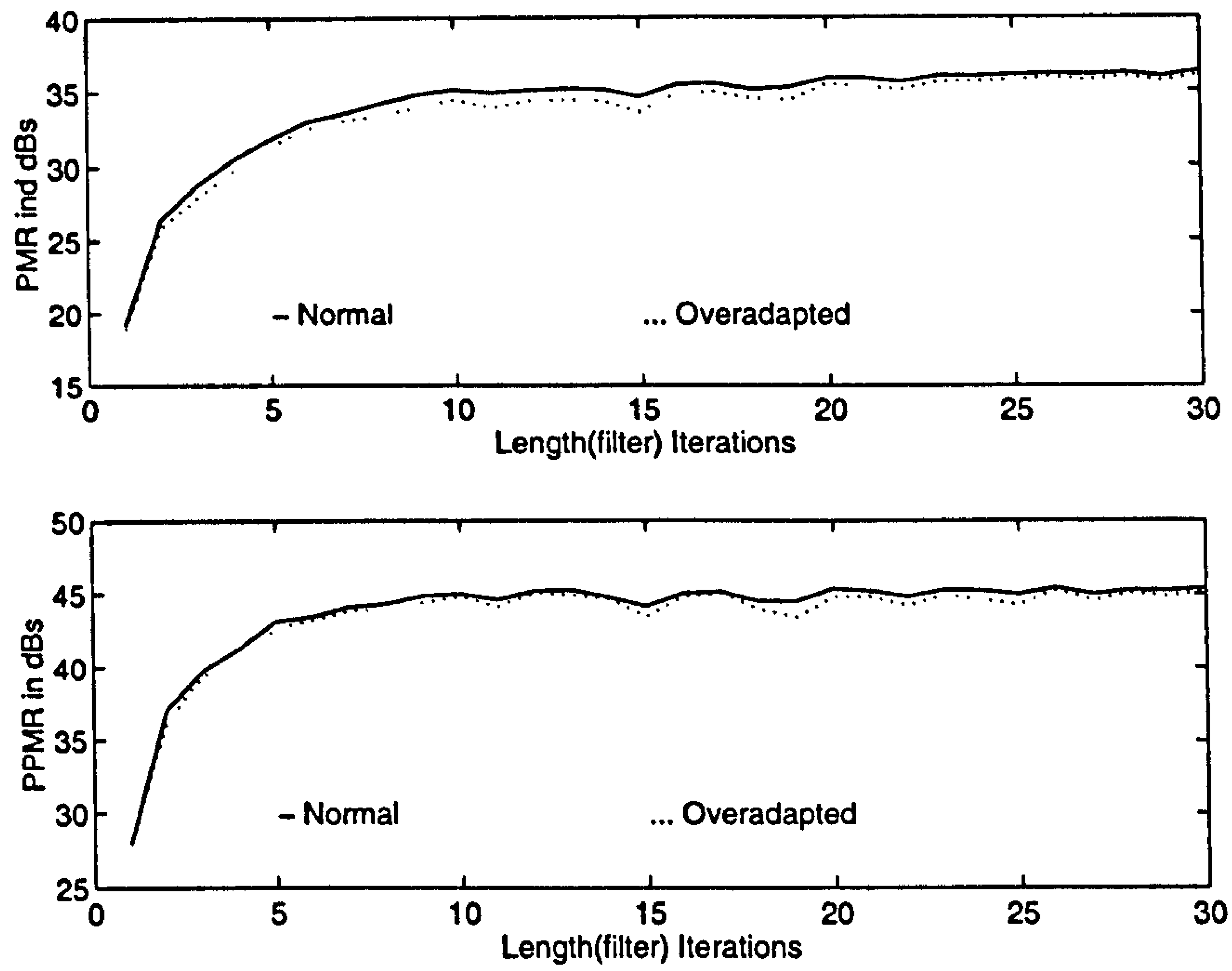


Comparing the results in terms of PMR and PPMR, the single channel system achieved 38.7dB and 45dB respectively but the two channel system produced figures of 72.8dB and 60.4dB. There can be little doubt of the advantages of utilizing the two stereo channels separately in the light of these results. The additional cost in processing power is modest compared to the improvement achieved, simply doubling the length of a single channel filter will require comparable processing power but without anything like the returns seen here. (See discussion in section 7.7.6.)

Over-adapting

It is possible to run the adaption loop of NLMS twice (or more) for each of the signal samples, this appears likely to yield a better convergence rate at the cost of additional processing power requirement.

Figure 7.19: PMR and PPMR for Standard and Over-adapting Filters



In practice the results (figure 7.19) do not show any useful improvement for the application of this technique.

7.6.2 Alternative and Hybrid Algorithms

There are a number of possible improvements that deviate further from the NLMS algorithm, these algorithms in pure form tend to require more information or significantly greater processing power but specialized forms can have advantages.

Hybrid Newton/NLMS Method

Newton's method is one of the basic adaptive processes but it is not commonly used as a basis for practical systems since knowledge of the input signal's statistics is required. Newton's method is generally faster to adapt, sometimes considerably so since it follows a direct path to the minimum rather than algorithms such as the NLMS which follows approximately the path of steepest descent. Unfortunately Newton's method is impractical, too much knowledge of the system is required and the processing requirement appears high. Widrow [62] suggested a hybrid LMS/Newton that appears to offer many of the advantages of the Newton algorithm without as severe limitations, he called it the Sequential Regression

Algorithm (SRA). This solution appears to be very similar to that suggested by Fujii *et al* [21].

More recently Petillion *et al* [49] published a paper that gives details of another version of the fast Newton method. This version is claimed to be able to offer the benefits of the Newton hybrid methods but requires only marginally greater processing power than the standard NLMS algorithm.

Frequency Domain Methods

Filtering can be carried out in the frequency domain, [8], [5] etc. Since the Fourier transform of a signal does not loose any information then there is no damage caused by Fourier transforming and the inverse Fourier transforming a signal. Borallo *et al* states "The large number of calculations required to implement a long FIR filter can be circumvented by either using a non-transversal structure or filtering in the frequency domain." He also states when talking about computational cost "... or an intolerable delay because of the data gathering needed for a frequency domain implementation." The algorithm he proposes is a partitioned frequency domain filtering system which effectively performs a time to frequency domain transform onto a non-linear frequency scale. The reduction in processing power required is obviously significant for relatively wide bandwidth signals.

It is possible to build equivalent adaptive processes in the frequency domain but there are few advantages. The processing requirements and convergence of the two methods are essentially similar. It was chosen to largely ignore these methods since they represent the fringe of adaptive technology and are less intuitive.

Recursive Methods

Recursive techniques are attractive since they are similar in form to a reverberant system and in many other applications are able to achieve comparable performance to FIR filters with fewer (sometimes far fewer) coefficients.

Typically recursive methods are not practical since the processing requirement for this type of algorithm is very much greater. For instance Hansler [28] indicates that the computational requirement for the Recursive Least Square (RLS) algorithm when used in the telephony problem is 10 length(filter) operations in each sample period as opposed to 2 length(filter) for NLMS [28]. Recursive algorithms also tend to be sensitive to computations error and sometimes other parameters.

There are fast recursive methods available, [53], but even these algorithms still require far more processing power than NLMS and after the reduced processing power requirement comes at the expense of robust stability.

Wavelet Based Methods

For wide bandwidth systems (several octaves at the very least) it may be worth considering the time/frequency resolution trades that are possible by making use of wavelet techniques. (The Fourier transform is a special case of the Wavelet transform.) In practice, in sampled data systems the application of wavelets translates to what has become known as sub-band filtering techniques. In this type of filter the signal is separated according to the frequency of its components and then different filters are applied to the various frequency bands. This allows LF filters of substantial length without requiring the massive number of taps that would have been needed if the full bandwidth were to be processed. (But if there is to be any saving requires that HF filters must be much shorter - a reasonable requirement for minimum phase signals but less applicable to acoustic systems.)

The application of this technique is common to the telephony echo cancelling problem and good results are reported [25],[65], [30].

It is uncertain if the technique would be able to offer very significant savings for acoustic inverse filtering, since the accuracy required is much greater. It is certain that the high frequency components of a signal are absorbed more readily, particularly by furnishings such as curtains and thus the high frequency signal will decay more rapidly. This alone is no guarantee that the filter required for these high frequencies will be shorter, though it would certainly be so for the telephony case.

There is little doubt that this method would have been tried if it had not been found that the required bandwidth is very limited and thus this technique has little to offer. (The reasons that the bandwidth required is limited are discussed in the next section 7.7.) There is one difficulty that would be experienced when applying this technique. The different length filters would have different optimum delays, since the delay is half of the filter length typically for optimum performance.

The optimum delay depends on the system, but generally it is of the order of half of the filter length plus the transit time. For LF filters it is sufficient to neglect the transit time since this is far shorter than the optimum delay. For the much shorter (in time) high bandwidth filters this will not be the case. The filters must always be roughly twice the transit time as an minimum [44], which will result in a significant number of coefficients at high frequency.

To ensure that the signals are synchronized when re-combined after filtering it is necessary to further delay the inputs to the filters with shorter inherent delay. This requires a delay line as long as the longest delay and at the sample rate of the highest bandwidth filter. The additional requirement does not increase the processing power significantly but it would require significant memory at full bandwidth.

7.7 Practical Limitations to the System

One area that seems to have received little or no consideration in the literature to date are the limitations of applying an inverse filter to a real acoustic system. Much of the research to date has featured very limited bandwidth and very short filters that were quite incapable of delivering anything approaching HiFi performance and thus the limited operation demonstrated in no way proves that the concept is practical. SigTech's inverse filter system [57] is slightly different, in that it is not adaptive, and thus it will not necessarily suffer the same problems as an adaptive filter.

A good indication of the potential performance of the inverse room acoustic filter can be achieved through basic simulations, though the results of direct measurement are also presented later in this section.

The limitations of the system will only be considered for stereo reproduction, though the results may be generalized with reasonable ease. The prefix stereo, from which the word stereo in audio terms is derived, comes from the Greek word stereos meaning "solid, having three dimensions." The phrase was originally coined since with careful recording and reproduction some sense of a virtual three dimensional image of the original sound may be perceived when using this system.

In section 3 the stereo reproduction system was analysed with respect to human hearing abilities. At the conclusion it was clear that, given the way the stereo system works, the human location system would probably not find sufficient location clues in signal components of less than about 1800Hz. (This figure is from analysis of the results of psychoacoustic studies and has not been directly demonstrated. Practical testing of this result is difficult since and non-reflective environment is required if only the limitations of stereo reproduction are to be investigated.) The only exception is for signals that occur from one channel only - but this is impossible for a stereo signal generated as Blumlein intended [54]. Currently, many stereo recordings are created by taking a separate signal for each instrument etc. and then assigning different amounts of every source signal to the two channels instead of the traditional crossed cardioid method etc. This new method will allow a source to be reproduced from one channel only, but if this were done the result should not truly be considered a stereo signal.

Binaural Reproduction

The simplest method of producing a "solid" image of a sound is to record exactly what each ear of the listener would have heard and then reproduce the signals recorded using headphones. If the recording microphones are located in a dummy head then even the complex interaction of the external structure of the ear (pinnae) can be reproduced in theory. In practice the very wide variation in the physical form of the pinnae makes this less effective than might be expected.

Binaural reproduction requires that the signal can be reproduced in each ear without crosstalk, unlike the stereo system which accepts that the signals from both loudspeakers will reach both ears. This means that it is not possible to use loudspeakers in a similar way to the stereo system to correctly reproduce a binaural signal. It is possible to alter the specification of the acoustic correction filter so that not only is each signal to be reproduced at the listener's appropriate ear without being corrupted with reflections but none of the signal will reach the listener's other ear. This approach has been suggested by Nelson *et al* eg. [44]. Thus for binaural application the filters have to achieve two separate constraints simultaneously. Technically this is achievable to a degree, though the additional constraint will probably reduce the effectiveness of the filter at cancelling reflections. Since it has already been indicated that achieving a sufficiently accurate filter is problematic the additional constraint required for binaural reproduction is probably impractical.

Stereophonic reproduction

The accepted method of producing an image of a sound-field was pioneered by Blumlein [54]. Instead of fighting the inherent problem in loudspeaker reproduction, namely that each of the listener's ears hears the sound from each loudspeaker, he designed a system that could tolerate it.

Blumlein's stereo system produces results that are generally held to be satisfactory, but only over a relatively small area. With this type of system it is desirable to optimize the path from the speakers to their closest ears, ie. remove the effects of reflections and effectively flatten the frequency response, but it is difficult to say exactly what the path to the further ear should be. The reflections should be cancelled for the the further ear to remove the effects of the room but the frequency response and additional time delay are an important part of the stereo systems operation and need to be preserved. In practice this is not an issue because of other, dominant, limitations that will be discussed later in this section.

Since most commercial recordings are designed for the stereo system this will be considered exclusively.

7.7.1 Main Problems with Acoustic Correction Filters

The simplest inverse filter system will optimize the path between the left hand speaker and the left ear and likewise for the right hand side ignoring the crosstalk signal paths. Any simplistic system targeted only at producing a perfect primary path runs the risk of corrupting the crosstalk signals, these signal paths are an intentional part of the operation of the stereo concept. But far more significant than these problems are the basic limitations the physical systems involved place on the concept. Even if only one reproduction channel were considered there are significant problems facing the proposed filter correction system:

- Distance between the sensor and the listener

  It is impossible to position the sensor at the same location as the listener, while it must be positioned as close as possible the distance between them will degrade the performance, particularly at the higher frequencies.

- Sensor accuracy

  The sensor must be a high accuracy type (known, compensatable performance defects are acceptable), since the filter will be adjusted according to sensor's perception of the acoustic signal. To match a good HiFi system a high grade microphone will be required, which will be very expensive. A byproduct of the NLMS algorithm is that the system should be tolerant to limited non-linear distortion. (Non-linearities will limit the ultimate accuracy that the filter can attain, but provided the levels are low will not significantly affect the operation of the algorithm.) The problem of producing a suitable microphone is substantially reduced if a reduced bandwidth can be accepted.

- Sensor Type

  For full bandwidth solutions some form of single point stereo microphone will be required, there are a number of slightly different alternatives available. In order to function properly these microphones should be positioned in free space. If a single-point stereo microphone is required there are a number of options that should be considered for possible practical advantages.

- Sensor shadowing

  Ideal sensor location will not be possible in a real system. Distance effects have already been mentioned but shadowing is also likely to be a problem. (It is expected that in a commercial system the sensor would have to be positioned behind the listener's head to be acceptable to the consumer. Current systems do not have a permanent sensor microphone, since they are not adapting in real time, and thus do not suffer from this problem.)

- Filter accuracy and convergence rates

  The final accuracy the filter achieves and the time it takes to converge are important considerations when designing a practical system. Many of the trade offs for filter performance are against processing power.

  Available processing power does constrain any systems being considered at the moment but given the continuing improvements in DSP hardware then provided that the requirements do not exceed the limits of current technology by many orders of magnitude suitable processors will become available in a reasonable time-frame.

- Background noise

  The effects of background noise on the realizable performance of the system also needs to be considered.

- Noise Pollution

  Applying the filter will increase the sound energy in the room for a given listening level and thus may cause additional problems of noise pollution. (Superposition correction inherently requires cancellation of the reflections of the cancellation signal which makes the problem more complicated and increases the sound energy present in the room.)

7.7.2 Distance between the sensor and the listener

Ideally the sensor microphone needs to be exactly at the listening point. In practice this of course will not be possible for a system that is required to operate continuously alongside the listener! For a marketable, convenient system the distance needs to be as large as possible. The minimum distance that might be achievable depends very much on the sensor type chosen but even so the distance is unlikely to be less than 0.1m and probably will have to be at least 0.2m in practice. For instance the sensor could clip onto the back-rest of the listener's chair at these distances.

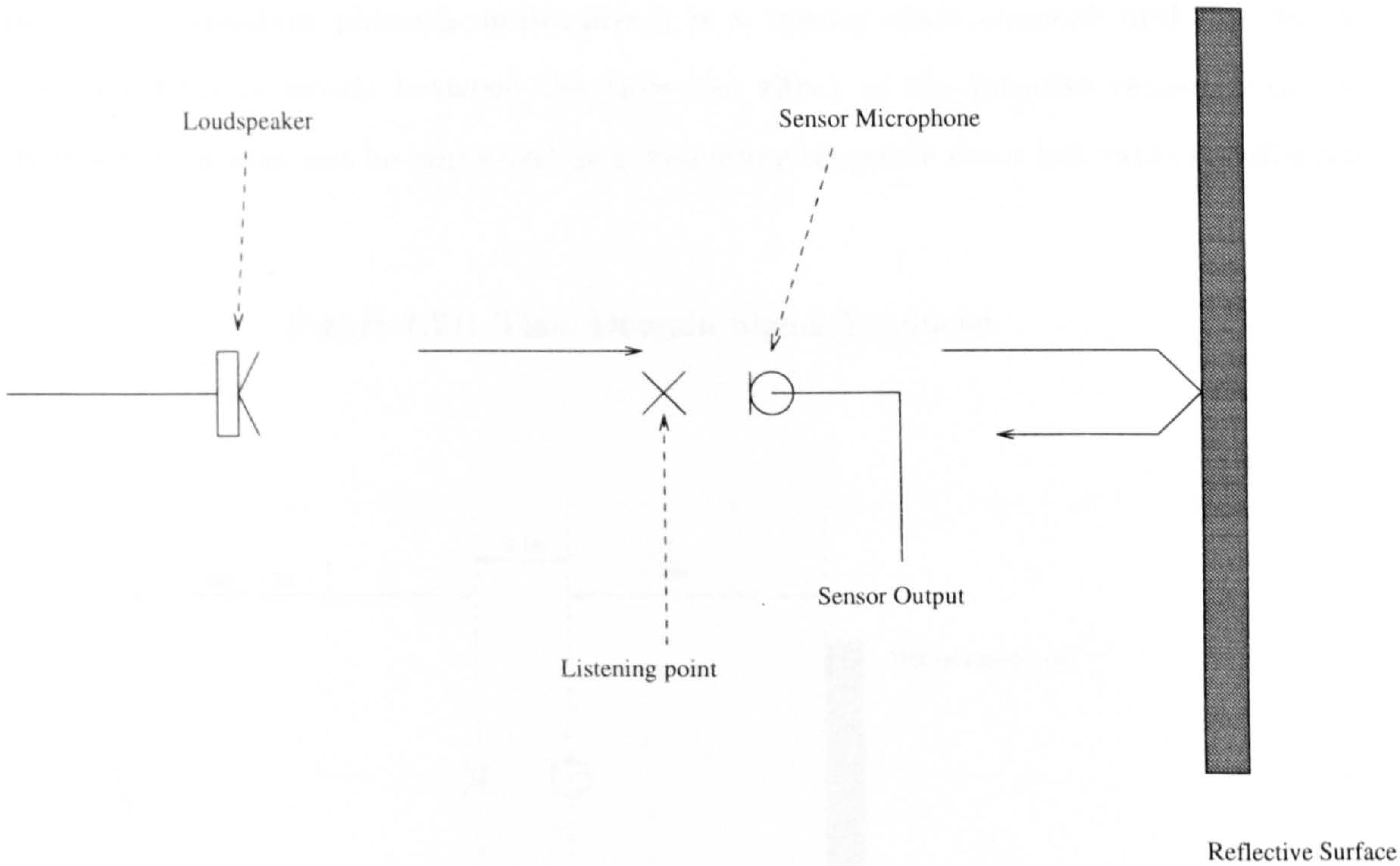
The effect of such a distance between the listening point on the impulse response can be estimated and even measured but it is difficult to make a judgement of quality on this basis. (It is possible to measure the impulse response that is obtained at the listening point with correction applied on the basis of the correction sensor.) Converting the measured impulse response to the frequency domain (by making use of the Fourier transform) the frequency response to the listening point can be obtained, which is easier to evaluate. (Deviations from the a perfectly flat frequency response are a well known and understood measure of audio reproduction quality.) The frequency response cannot be used with impunity, however, since the impulse response is long; the ear may well not perceive the effects as limited to amplitude/frequency aberrations but reverberation or even echoes. Note that this problem only exists due to the transient nature of music.

Model Based Analysis

The simplest model for a reverberant room that can be conceived is an acoustic transmission path with a single reflective surface, as shown in figure 7.20. Using this simple one dimensional model it is possible to predict the signal received at the listening point and at the sensor. More complicated models could be used but the effects of the reflector become difficult to predict.

The problem of the distance between the listening point and the sensor, when considered in a single dimension, can be expressed as an acoustic source in front of and a reflector behind the listener. The effect of distance between the sensor and the listener can be calculated. (Note that no back wall is included since as the loudspeaker is considered to emit sound forward only and the added complications of considering the reflections of reflections is ignored at this stage.) The figure 7.20 shows this simplified setup:

Figure 7.20: Model Layout



If an impulse were produced by the loudspeaker then assuming that the filter had already adapted such that there was no error at the sensor a second impulse would be produced to cancel the first, and then a third to cancel the reflection of the second etc. Given a perfect system the cancellation would be absolute at the sensor but the listener, being removed be a small distance would be presented with two opposite and slightly separated impulses. If the listener were closer to the speaker than the sensor then the correction impulse would arrive before the reflection and if further then later. The frequency response of this impulse pair could give useful indications of the limits of the system with increasing separation between the listener and the sensor.

For the purpose of demonstrating the scale of the problem it is necessary to make a number of assumptions, the reflectivity of the wall is assumed to be 50%, constant at all frequencies and the layout is assumed of the dimensions given in the diagram. The sound pressure, which should have been inversely proportional to r^2, is initially assumed to remain constant while the signal is being transmitted through the air. (r is the nominal distance from the loudspeaker. A small correction being required otherwise a singularity would result at $r = 0$, the easiest option is to take $r = .15m$ at the loudspeaker, which corresponds to a typical cone radius of 20cm.)

Figure 7.21 gives a much clearer illustration of the waveform produced. The time domain representation of the signal at the listening point can now be plotted and this then may be readily transformed to give the frequency response. It should be noted that the frequency response (unless absolute phase is maintained) is a steady state concept and due to the relatively long time intervals between the taps the effect of the impulse response on the sound reproduction may not be perceived as a frequency response error but rather as discrete echoes.

Figure 7.21: Time Domain Signal Produced

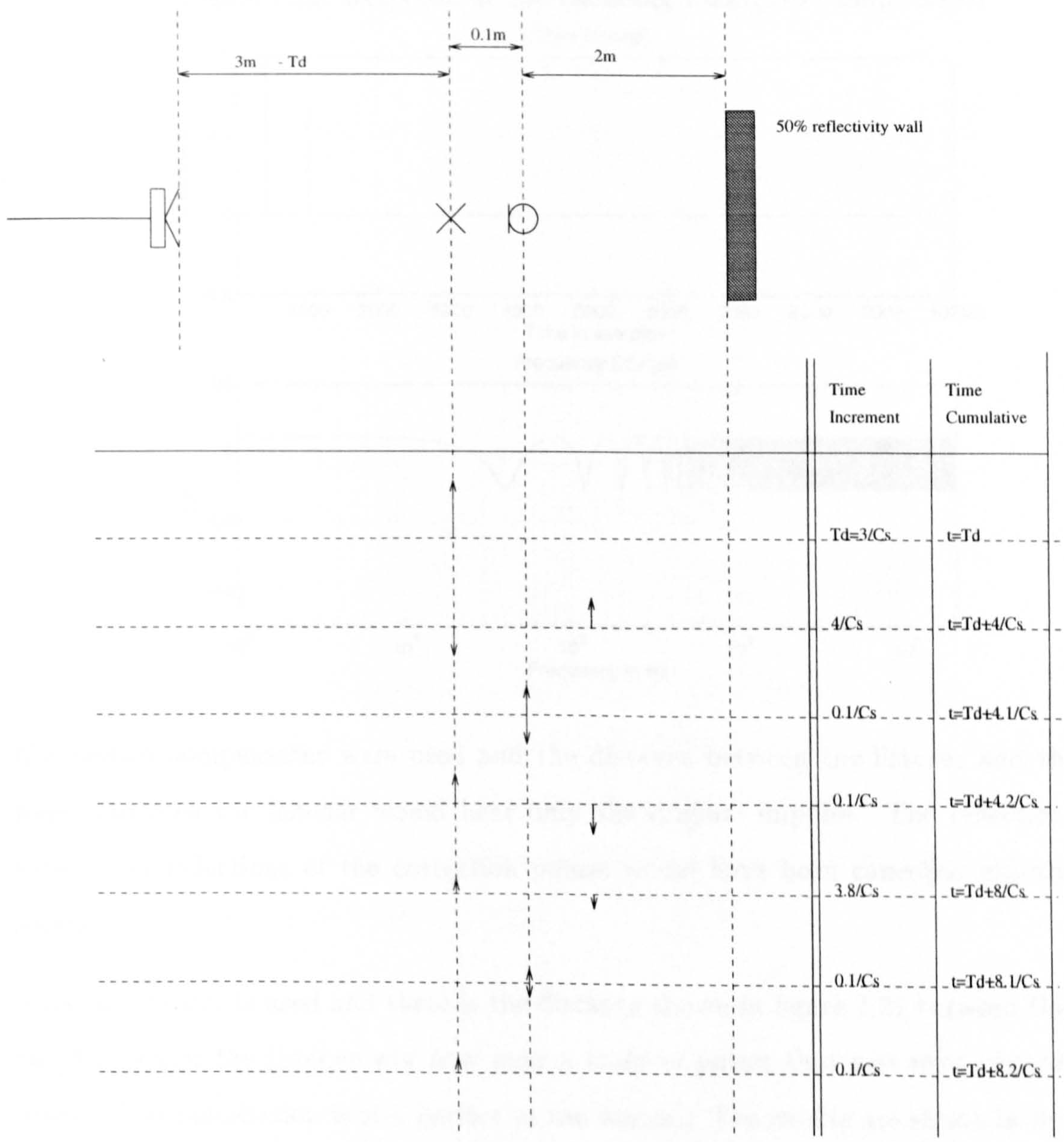
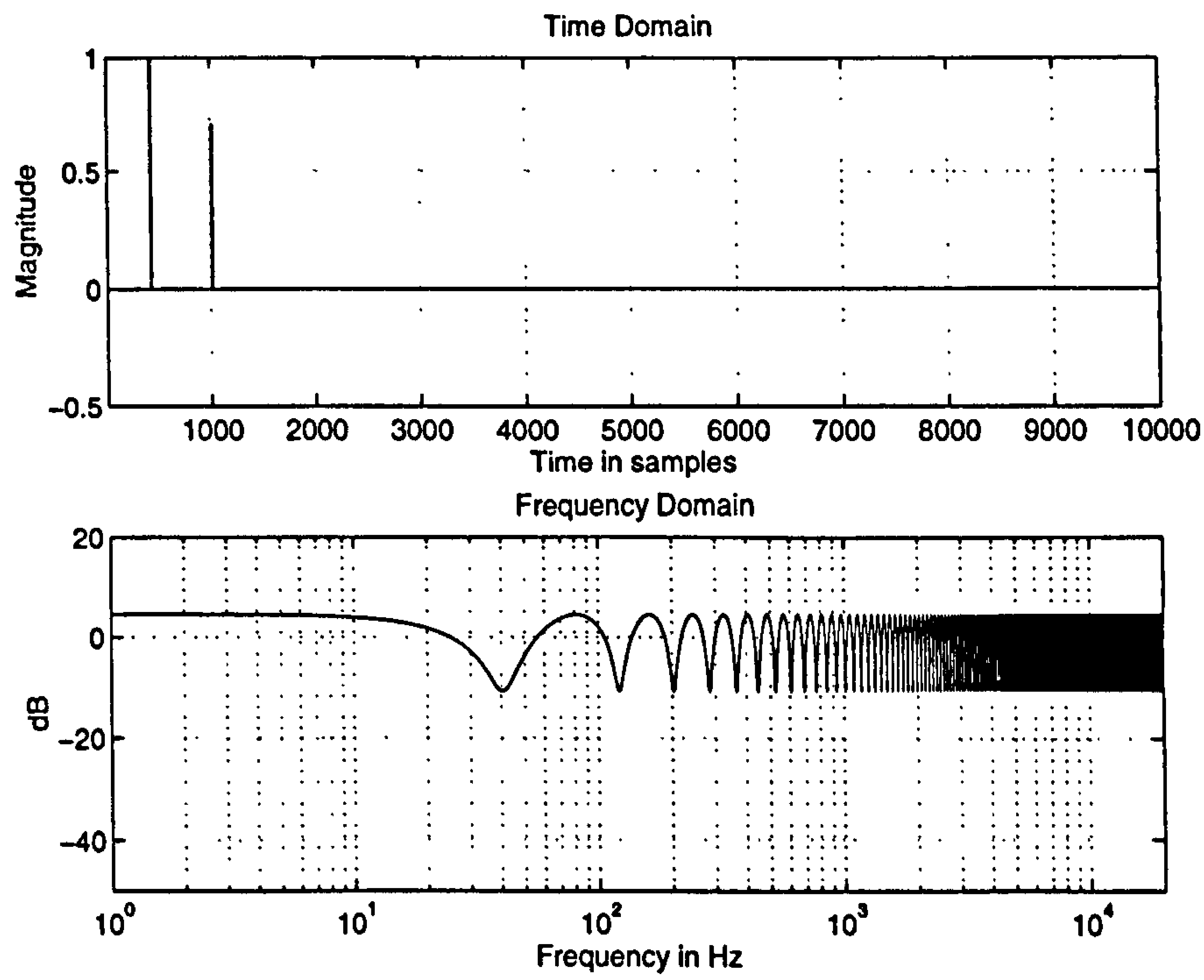


Figure 7.21 shows the spatial location of the impulse signal (across the page) against time (down the page). The location of the impulses is shown as a snap-shot at certain interesting points in time. Thus the original impulse reaches the listener (represented by X) at time

$3/C_s$ where C_s is the speed of sound. This time interval is assigned the label T_d for convenience in the cumulative time column. It can be seen that the reflected impulse is exactly cancelled when it reaches the sensor at $T_d + 4.1/C_s$ cumulative time.

In a system that had no compensator the signal would simply be reflected from the back wall once. Thus the listener would hear only the original and one reflected version. This situation is shown in figure 7.22. The time domain plot is what the listener would hear if an impulse were transmitted and the frequency domain plot is the spectrum of this signal. The latter is also the effectively frequency response of the loudspeaker/listener path.

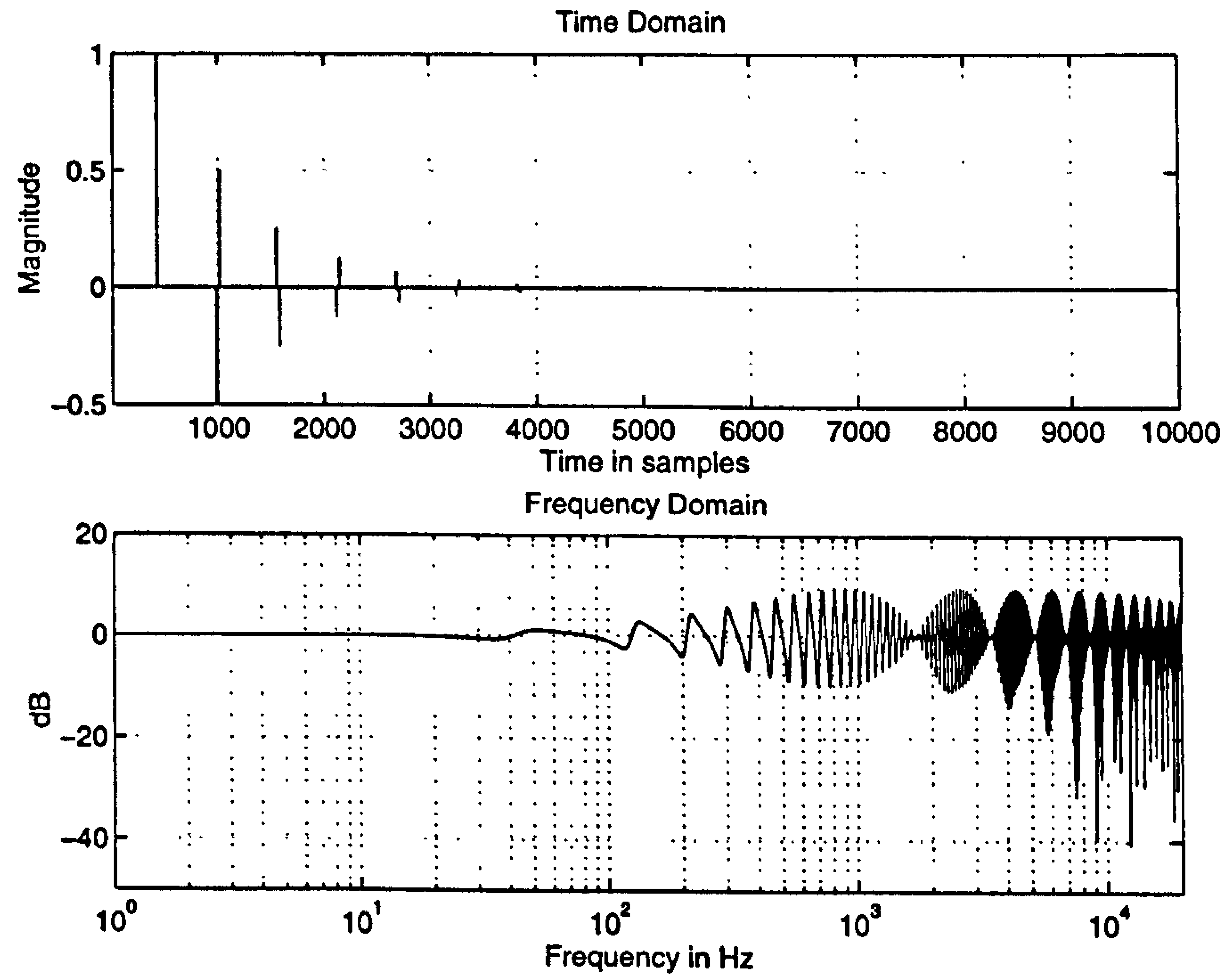
Figure 7.22: Response at the Listening Point, No Compensator



If a perfect compensator were used and the distance between the listener and the sensor were zero then the listener would hear only the original impulse. The reflection and all subsequent reflections of the correction pulses would have been cancelled exactly at this location.

If a compensator is used and there is the distance shown in figure 7.21 between the listener and the sensor the listener will now hear a train of pulses that just miss cancelling each other. (The cancellation is still perfect at the sensor.) The results are shown in figure 7.23. The correction achieved is far from perfect (which would have been a single impulse or a completely flat frequency response) but represents an improvement on the un filtered room at low frequencies. At high frequencies the results are rather worse than the unfiltered room. The break-even point is of the order of a few hundred Hertz because at higher frequencies than this the noticeably higher order deficiencies in the compensated response will probably be found to be subjectively poorer. (The rapid changes in phase associated with high order systems are known to be subjectively poorer.)

Figure 7.23: Response at the Listening Point, Compensator used



The model of the room system used at this point is highly simplified, some of the assumptions that need to be addressed are:

- Reflectivity

  The walls have been assumed to reflect all of the frequencies evenly, this is unlikely to be the case. If the wall were to be made of hard plaster then the results would be similar to the initial assumption except that the reflectivity of a plaster wall is much higher across all frequencies. Carpets, curtains and other soft furnishings have much more interesting absorption profiles.

- Perfect filter

  To evaluate the concept it is reasonable to assume that the filter is perfect.

- Distance effects

  It is difficult to know how best to model the effects of distance on the SPL, in an open site the SPL will be inversely proportional to the square of the distance. This does not represent energy lost, rather it is due to the increasing surface area of the wave front. In a real room a large proportion of this energy would be reflected since it is a complete enclosure. On balance it seems that there is little point in accounting for the dispersion that would occur in a real system since this additional complexity would only trade one inaccuracy for another. If the model were expanded to include more reflective surfaces this tradeoff should be reconsidered but additional reflectors would substantially increase the computational complexity of the model.

- Number of surfaces

  At the moment only a single wall has been considered, obviously a dramatic simplification. Some indication of the effect of this simplifying assumption may be gained by introducing a second surface again on axis with the speaker but this time behind it. It is not realistic to develop a model that includes all six surfaces of the room by extension of this intuitive method since the reflection paths become far more complicated. (The problem is tractable if finite element analysis techniques are used but this becomes highly computationally intensive without yielding very accurate results.) The model used at the moment should be a worst case approximation since the distance between the listener and the sensor in the direction of the wave front will reduced for sound reflected from he side walls and the floor/ceiling.

While some of these limitations of the model may be addressed in the final analysis it will be necessary to make real measurements to establish the actual impact of the distance between the sensor and the listener. It is possible to extend the model relatively easily to include more realistic reflective surfaces.

Reflectivity is usually measured in terms of absorption, thus a surface with a 60% absorption coefficient surface will absorb 60% of the energy. "Sound absorption coefficient (.. also called the acoustic absorpity) is defined as the fraction of the incident sound energy absorbed by a surface or medium, the surface being considered of infinite area. - International Dictionary of Physics and Electronics - Vann Nostrand" The measure obviously has no units by definition.

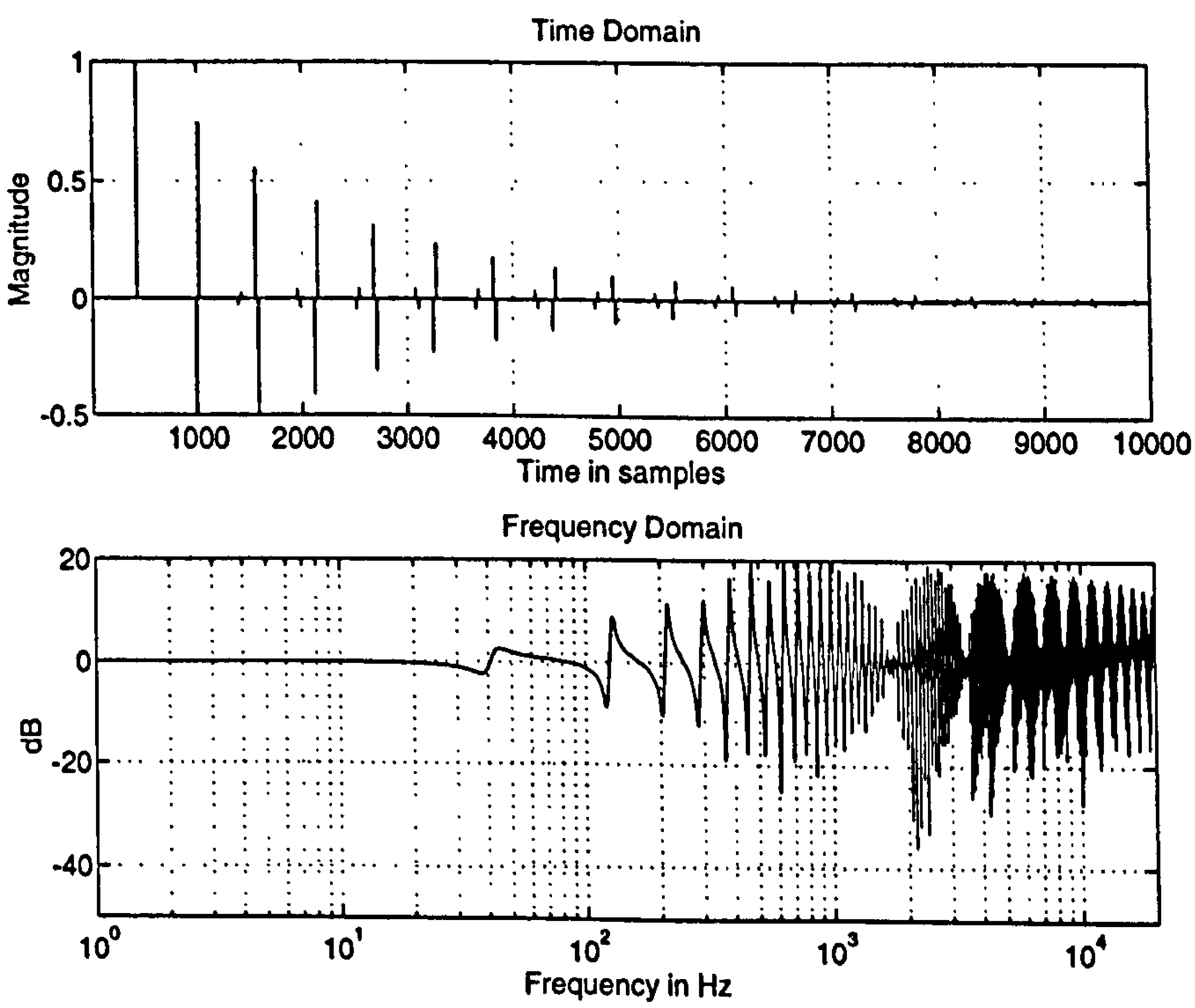
The data presented here is all drawn from "Audio Electronics Reference Book" [54]

Absorption Coefficient (0 is perfectly reflective)						
Material	Frequency Points					
of Surface	125	250	500	1k	2k	4k
Carpet	0.07	0.25	0.5	0.5	0.6	0.65
Hard Plaster	0.03	0.03	0.02	0.03	0.04	0.05
Curtains	0.05	0.15	0.35	0.55	0.65	0.65

If this data is extrapolated in a suitable manner then it is possible to incorporate it in the simple model. The extrapolated data points are not critical to the modelling process since they fall mostly outside the band of interest.

If the reflective surface is chosen to be hard plaster then the resulting time response is highly reverberant and all frequencies above about 100Hz are seriously affected. Fortunately most real rooms contain a significant amount of soft furnishings so this model is unreasonably extreme. (Note that the "blips" between the normal reflections and correction pulses are spurious. The cause is related to the method of using the absorption/frequency data. The small values of these blips should not affect the data significantly.)

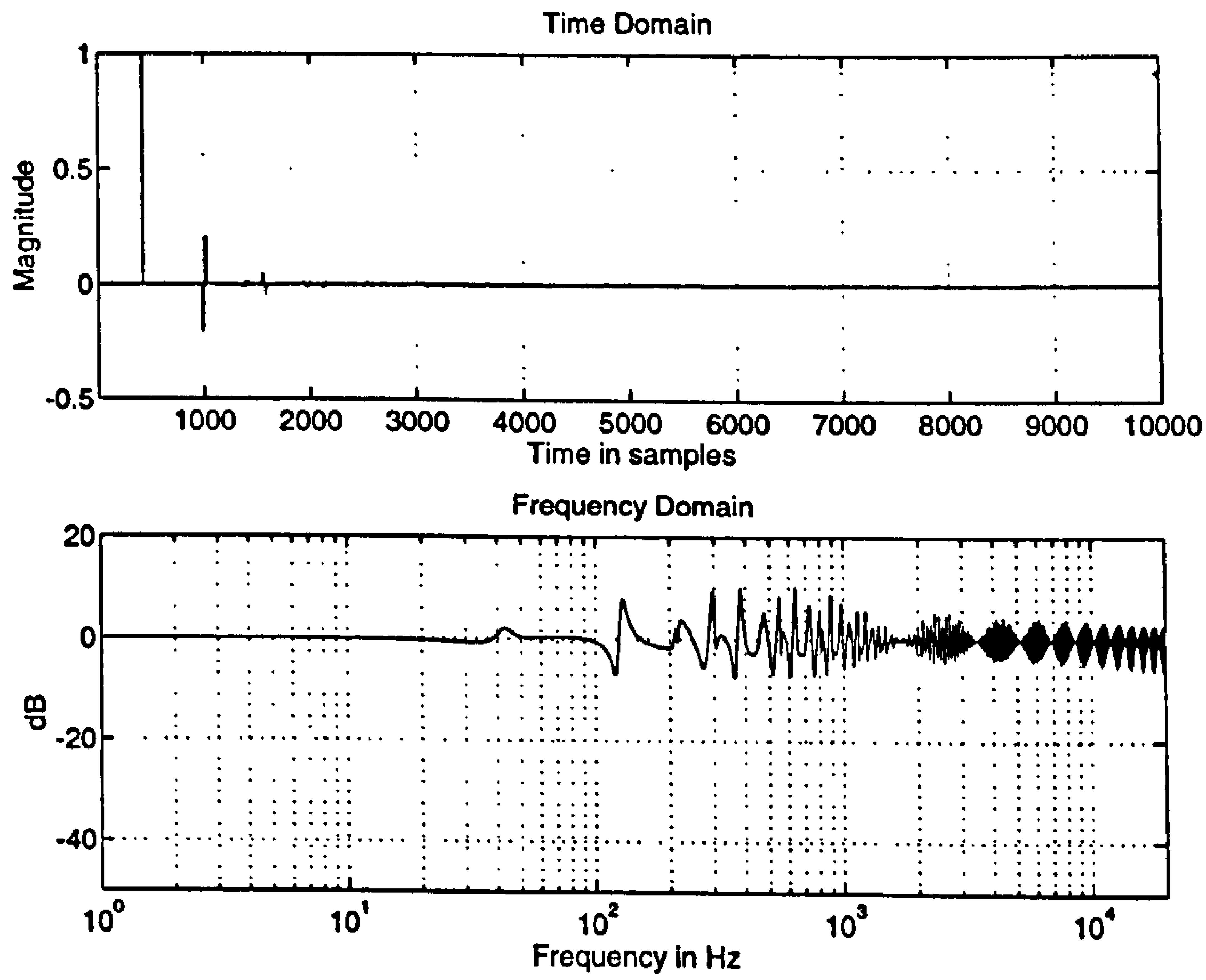
Figure 7.24: Model Response - Plaster Walls



A more useful model can be made using curtains as the reflective surface, since this is probably a better approximation to a real room's mix of hard plaster walls and soft furnishings.

The results are quite different but at LF the first significant errors still occur at just over 100Hz. The results are predictable since curtains will have almost no effect of low frequency signals, thus they will be reflected by the hard surfaces behind, but there will be much more substantial attenuation of mid and high frequencies.

Figure 7.25: Model Response - Curtains



To summarize, it appears that it will not be possible to apply a correction signal effectively at frequencies significantly greater than about 100Hz due solely to the physical separation between the listener and the sensor.

Real Measurements

The models indicate that for separations between the listener and the sensor as low as 10cm the compensation concept is only viable below a few hundred Hertz. Unfortunately many assumptions had to be made and thus it is difficult to place a bounds on the accuracy of the results. The next logical stage is to attempt to measure a real system to confirm these preliminary results.

There are a number of different methods that could be utilized that use varying amounts of real information. The first two methods have the advantage that they can be carried out with reduced bandwidth if appropriate.

- Minimal Measurements

  If the impulse response is measured at two locations separated by the distance in question (using PRBS sequences as discussed earlier) and then a filter is developed for one of these locations (using whichever of the methods discussed seems most appropriate) the detrimental effects of the distance can be simulated using by using convolution to find each system response.

- Real Degradation

  It is possible to take the two impulse responses as measured before and develop a filter for one of them but instead of simulating the effectiveness (by convolving to give the system impulse response) the response of the system can be measured when enhanced by the filter at both locations. This method is much less convenient due to the time required to make the additional measurements. An additional problem is that the impulse response for any real room will change over time and thus a filter will not remain valid.

- Subjective Testing

  It is theoretically within the capabilities of the test system to produce test signals for subjective comparison. The room impulse response could be measured at a location and then an inverse filter generated. Applying this filter to a music signal should result in near perfect reproduction at the point where the room impulse response was measured with degradation due to distance apparent as the listener moves away from this ideal. There are many disadvantages to this method, most importantly filtering the music signal without specialized hardware is a slow process but one that needs to be accomplished quickly, while the filter is reactively valid. Another problem is that moving the listener in the sound field will have a significant impact on the room impulse response that cannot be conveniently compensated for. There are also many problems associated with quantifying subjective tests and concentrating them on only the impact of this concept. (The test system is a long way short of HiFi quality and thus relatively subtle effects are difficult to discern.

In practice the first of the methods discussed was used for its simplicity and the consistency of the results generated. Some additional steps were included in the process to reduce the effects of certain errors.

- Measure a room impulse response
- Build and test a filter for this impulse response
- Re-measure the room impulse response (see note)
- Move the sensor a small distance, e.g. 10cm
- Measure the altered room impulse response
- Try the filter on the new impulse response

  (note) The room impulse response was re-measured so that if errors occurred in the measuring process then similar ones to those that occurred in the altered room impulse response would occur in this control. All comparative tests use this second version of the room impulse response.

There is no need to carry out the test for the full bandwidth since the filter is known to be unlikely to be effective. If this assumption is incorrect then the filter will prove good for the entire bandwidth measured. Reducing the bandwidth required makes the calculation and measurement problem far easier to deal with.

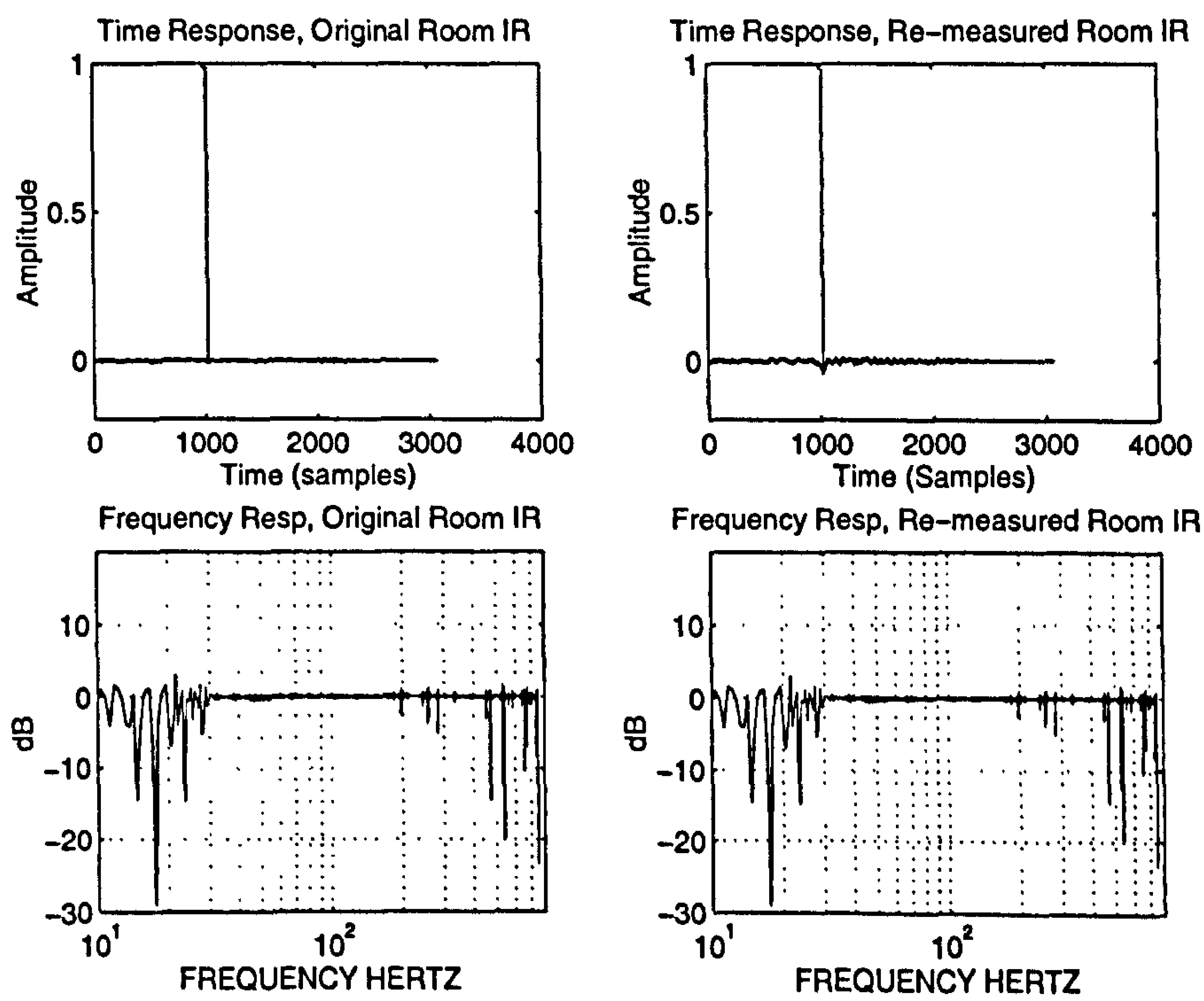
Due to the nature of the test the loudspeaker is considered part of the room system and the inverse filter will attempt to correct for linear deficiencies in its response. It should be noted that the basic unit used for this test suffers from some non-linear distortion which will tend to corrupt the results obtained here.

Experimental parameters were chosen to give a sample rate of approximately 1500Hz, which equates to a bandwidth of approximately 750Hz, which is adequate since the limit is expected to be of the order of 100Hz from the previous experiments. The reduction in bandwidth is convenient when developing high grade filters. (The size of the filter is limited by the computational facilities if the convenient matrix method is to be used.) The "matrix inversion" method was used to calculate the filter since this yielded the most accurate results in a reasonable (less than an hour) length of time. Due to the limited size of the problem the DEC-alpha platform proved sufficiently powerful to deal with the calculation. The filter developed was twice the length of the impulse response measured of the room to increase the accuracy of the results obtained - a total length of 2048 points.

The room impulse response was re-measured in an attempt to establish the error limits of the measuring system. The room used for this experiment is situated of the ground floor of the electronics main building and is therefore subjected to quite high levels of low frequency noise, both continuously from the heating system etc and sporadically from doors closing and movements of people on the floor above. In practice this results in considerable error at frequencies of less than about 20Hz.

Figure 7.26 shows the time and frequency domain plots of the system with both the original version and re-measured version of the room impulse response at the reference location. Both responses are filtered by the correction filter. It can be seen that while there are differences between the responses they are mostly limited to the extreme LF, their cause is probably the previously noted LF noise. The time response plots are the linear amplitude against the time in samples, where the simple frequency is approximately 1500Hz as stated previously.

Figure 7.26: The Filtered Response of a Real Room at the Reference Location



Considering figure 7.26. The room impulse response was measured and an inverse filter created for this impulse response. The system response, that is the response of the filter and room convolved, is the left hand plot labelled "Time Response". Underneath this plot is a plot of the Fourier Transform of this system impulse response. The room impulse response was then re-measured, in an attempt to establish some idea of the variability. This response was convolved with the filter developed for the previous measurement and the resulting system impulse response is given in the upper right-hand plot. Again this result was Fourier Transformed to give the plot below. Even the re-measured version of the room impulse response was filtered accurately by the filter from between about 30Hz and 400Hz. (The difference between the plots is small over the entire range shown.) The similarity between the two sets of plots in figure 7.26 demonstrates that the errors due to measurement are small over the band of interest and thus any effects observed when the microphone is moved are likely to genuine products of this alteration to the system.

Measurements of the room impulse response were taken at 10, 20, 40cm displacement on axis. These impulse responses were then filtered with the filter developed for the initial measurement. As expected the value of the correction filter becomes steadily more limited as the distance increases. If the distance is less than roughly 0.2m the quality of the system could be improved with the application of a suitable filter for frequencies up to about 100Hz. As has already been pointed out the FFT of the response is only an accurate measure for continuous signals and thus the impulse responses are also reproduced here, however, as has already been indicated it is difficult to make useful observations from this type of plot. It is possible to calculate the PMR of the system impulse responses for various displacements but this figure is a performance measure for the entire bandwidth only and thus would not be useful in estimating the useful bandwidth of the filter system.

Figure 7.27: Time Domain Responses of Systems at Various Displacements

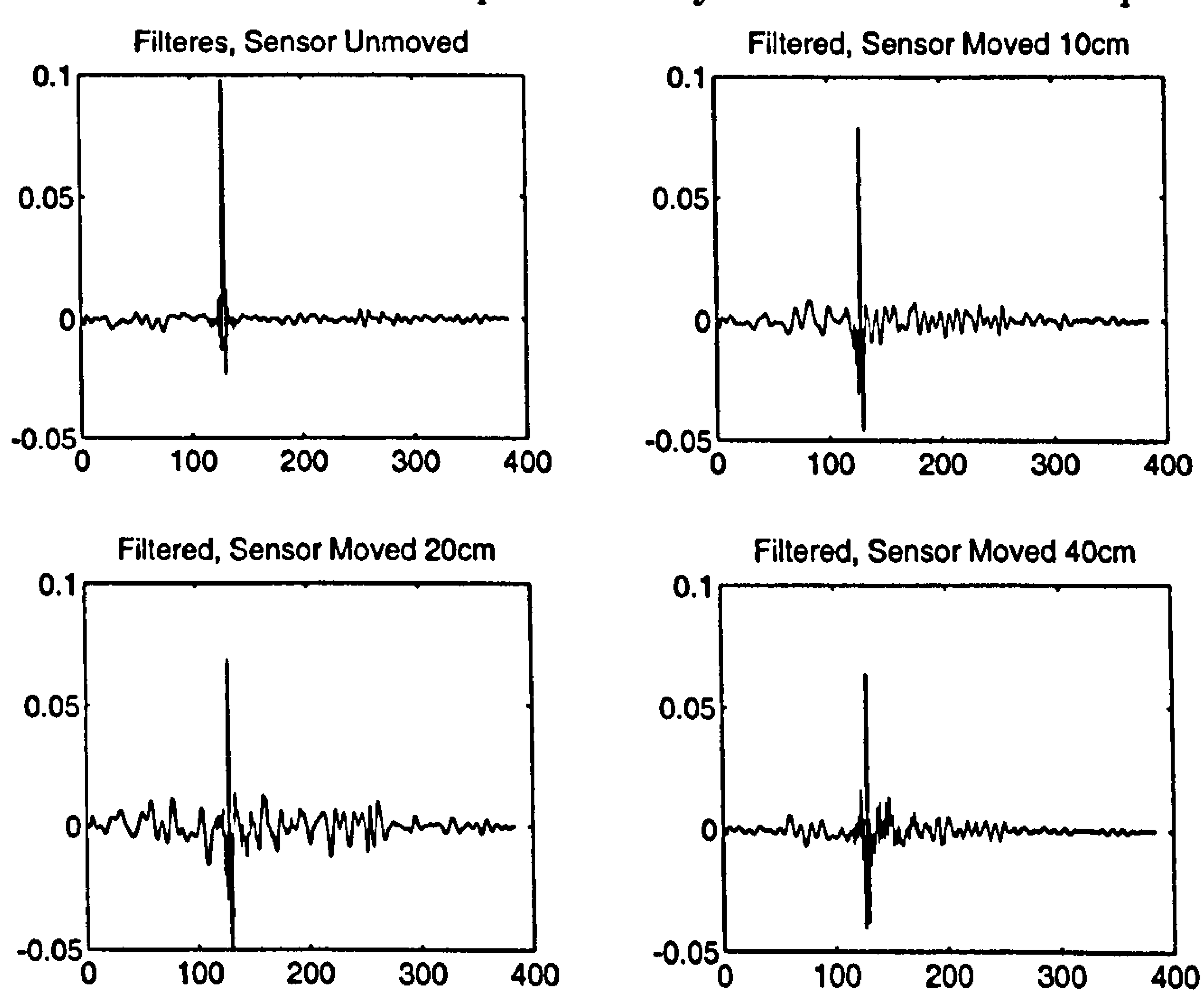
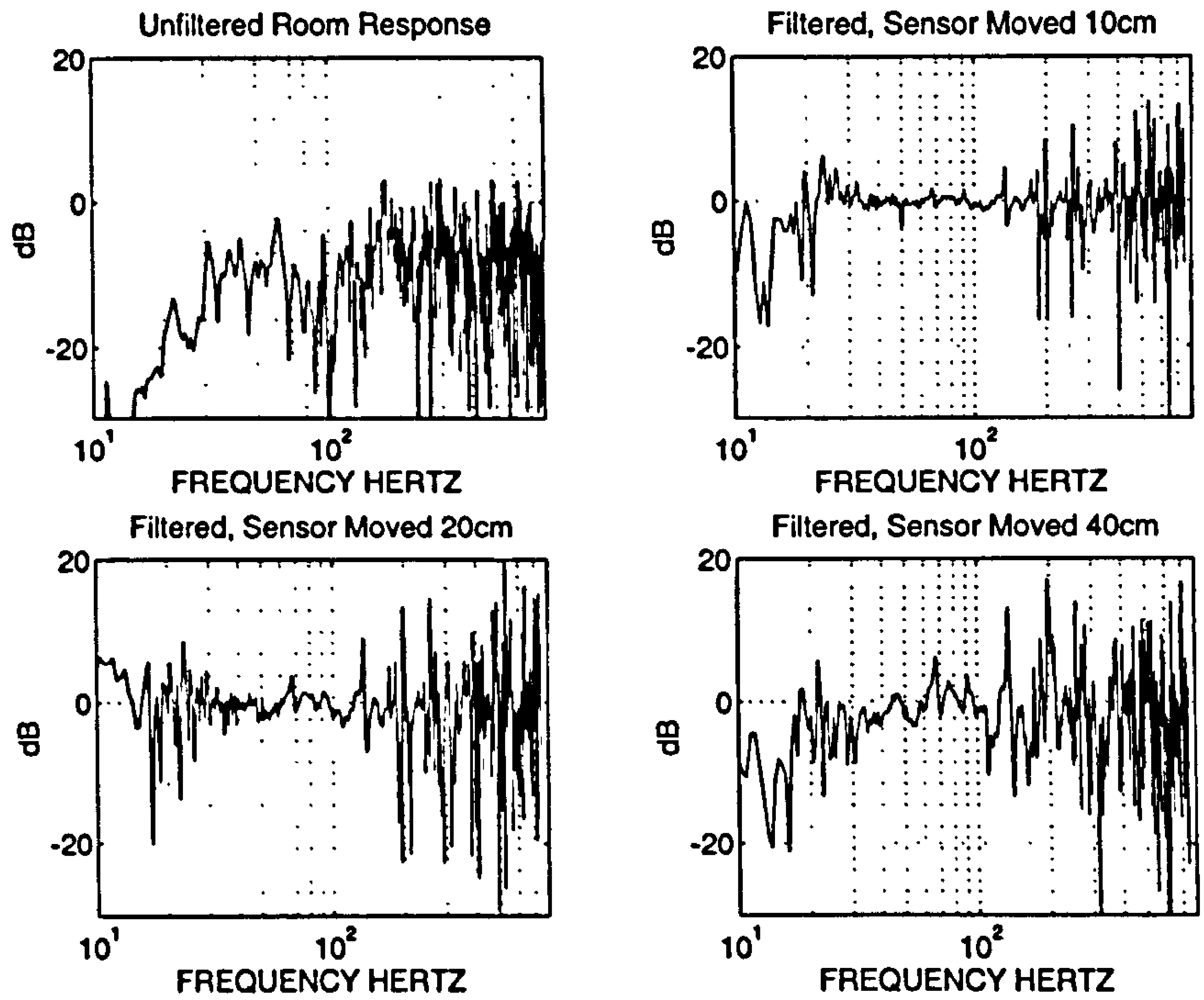
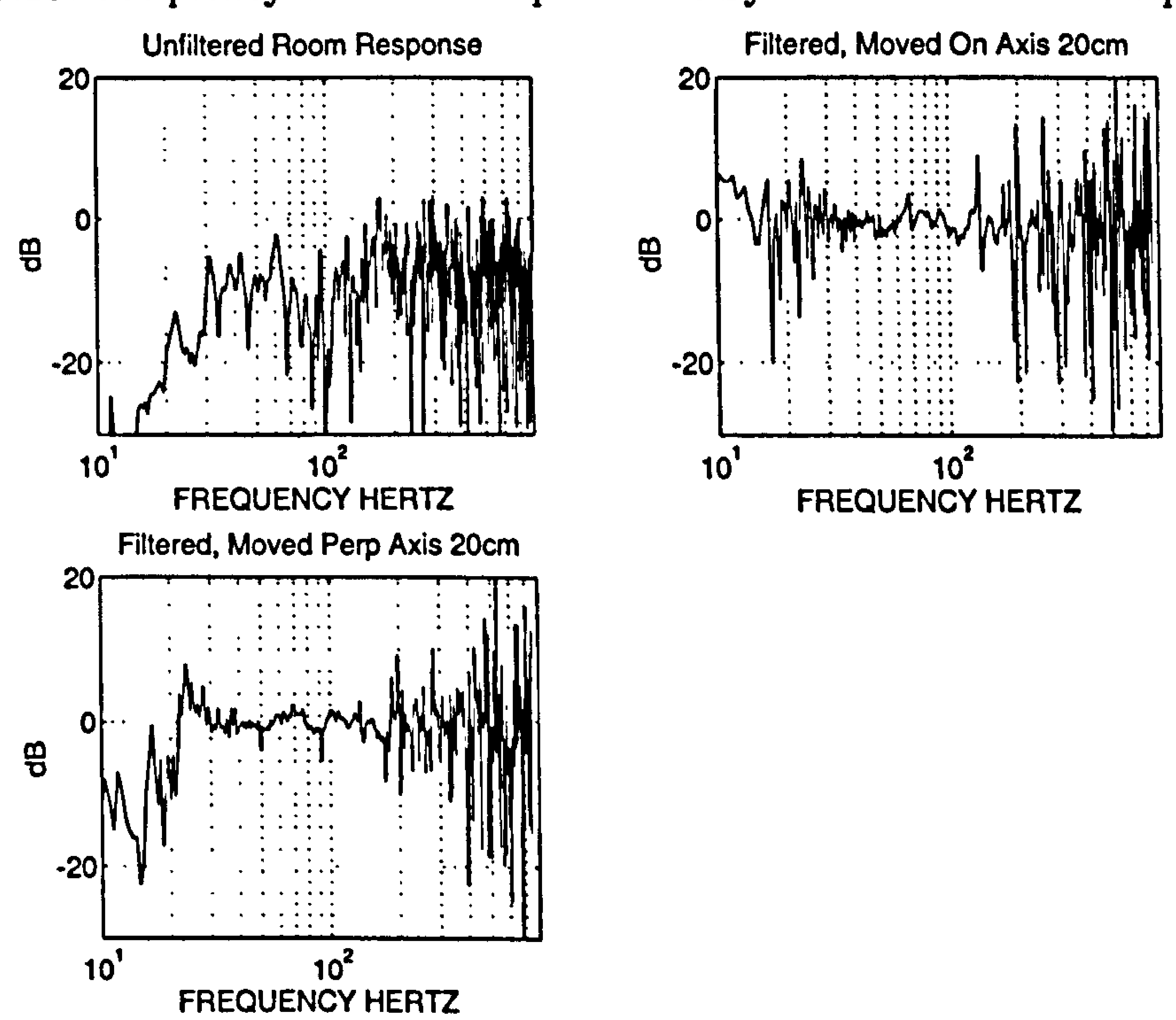


Figure 7.28: Frequency Domain Responses of Systems at Various Displacements



It is also interesting to consider the effects of displacing the microphone perpendicular to the axis of sound travel. It is expected that the effects will be less severe, which is born out in practice. Unfortunately it is difficult to imagine a system that would be acceptable to the consumer that had the sensor displaced to the side of the listening position.

Figure 7.29: Frequency Domain Responses of Systems at Various Displacements



To conclude; both the model and the real measurements have indicated that the utility of the proposed system is severely constrained by the inability to locate the adaption sensor exactly at the listening point. As expected the model results indicate the problem to be more severe than the equivalent real measurements however, the proposed model is sufficiently accurate to provide a useful first approximation. Disregarding any limitations other than distance between the sensor and the listening point then it is reasonable to suggest that a useful system could be made operating on the components of the audio signal below 100Hz.

It is possible to sub-sample the audio data and thus make savings in the processing power required provided that there is an appropriately delayed bypass route for the remainder of the bandwidth. It is necessary that the outputs of the sensors are band-limited in the same way as the input signal. This is only true since the acoustic system is largely linear.

The accuracy of the microphone and loudspeaker have no effect on these tests. Since neither of these element's responses are expected to change significantly with location, deficiencies in their responses will not affect the results and since the conclusions are drawn by comparison of the results at different locations there is no requirement for absolute accuracy. In addition the responses include a filter designed to compensate for deficiencies in either element.

7.7.3 Sensor Accuracy

The accuracy of the sensor limits the accuracy of the system. Given an unlimited budget there should be no difficulty in obtaining a sensor of sufficient accuracy and thus this does not represent an absolute constraint. Given that the distance between the sensor and the listening point alone will limit the useful frequency response over which correction may be considered to less than 100Hz the problem becomes much easier to deal with; the response of the sensor is only required to be flat over a single decade.

7.7.4 Sensor Type

Referring to the introduction to the human auditory system and the stereo system it is clear that there is little stereo information in the signal at frequencies less than about 1800Hz that the human location abilities would be able to use. It has already been found that the limit at which inverse filters are useful will not be greater than about 100Hz and thus there is no need to use any form of stereo microphone technique - a single microphone will be adequate.

It is generally accepted that there is little or no stereo information at the extreme LF of the audio range and therefore there are unlikely to be marketing problems associated with only using a single simple microphone. (For this reason there is only one LF channel in the DVD specification etc.) This neatly side-steps choosing a suitable single point stereo microphone and reduces the cost of the sensor dramatically.

A further advantage to only using a single sensor is that there are no crosstalk problems to consider and thus the difficulty in deciding how to handle the crosstalk path with the stereo system is not an issue.

7.7.5 Shadowing

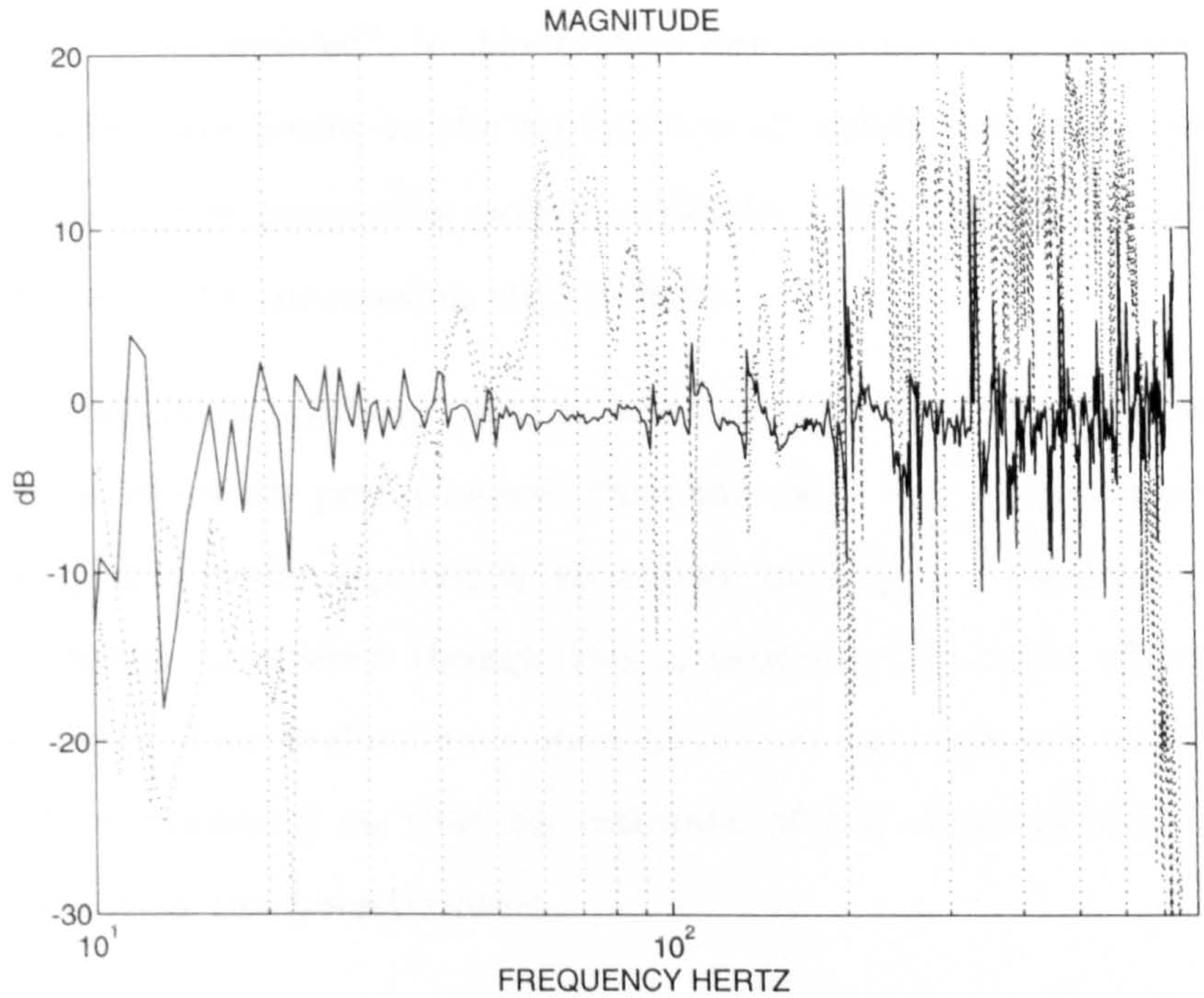
Since the sensor is required to be located close to the listener there will inherently be problems with the listener shadowing the microphone. Providing that there is a clear line of sight between the microphone and the loudspeaker then the effects will probably be limited; for instance mounting the microphone above the user's head.

Diffraction effects also are only significant when the size of the object is of the same order as the wavelength. Given that the human head is of the order of 0.3m in diameter then this effect is likely to be significant for frequencies of about 1kHz.

It is generally accepted that the shadowing effects are small provided that the object in question is smaller that about one wavelength of the signal. At 100Hz (the approximate limit of useful application of the inverse filter system) the wavelength is of the order of 3 meters, substantially larger then the human head. Provided that the approximations are valid then the large margins indicate that the sensor could conveniently be located on the back of the listener's chair without significant shadowing.

In practice, while a model could be created to produce a more accurate estimate, the easiest way to check the result is to measure the effect directly. The room impulse response is first measured with microphone unshadowed and then with a listener's head positioned directly in front of the microphone. It should be noted that the small movement of the listener to allow this would result in some change to the room impulse response. By comparing the differences between the response with the listener present and not present the effect of shadowing should be easily determined.

Figure 7.30: Comparison of Frequency Response with and Without Shadowing



Solid line: Difference between shadowed and unshadowed responses.

Dotted lines: Shadowed and unshadowed responses.

The plot shows clearly that there is very little shadowing effect for frequencies of less that 100Hz. The differences at the very low frequency end stem from the effects of noise, as before, it is more prominent here since a lower number of repeats were used when the responses were taken with a human head shadowing the microphone due to time required.

7.7.6 Filter Accuracy and Convergence Limits

As noted in the section discussing methods of measuring filter quality, if the aim were to build an adaptive filter to HiFi accuracy for the full bandwidth then it would be very difficult to achieve sufficient accuracy with a sensible processing power requirement. Fortunately, since it has already been demonstrated that the concept would be ineffective at frequencies of greater than about 100Hz; there is no need to deal with the full HiFi bandwidth. The effective sampling frequency is reduced by more than two orders of magnitude.

The optimum filter can be calculated from the original expression of the NLMS problem, the details are considered in section 2 “An Introduction to Adaptive Systems”.

$$\mathbf{h} = \left[\mathbf{G}'\mathbf{G}\right]^{-1}\mathbf{G}'\mathbf{d} \tag{7.6}$$

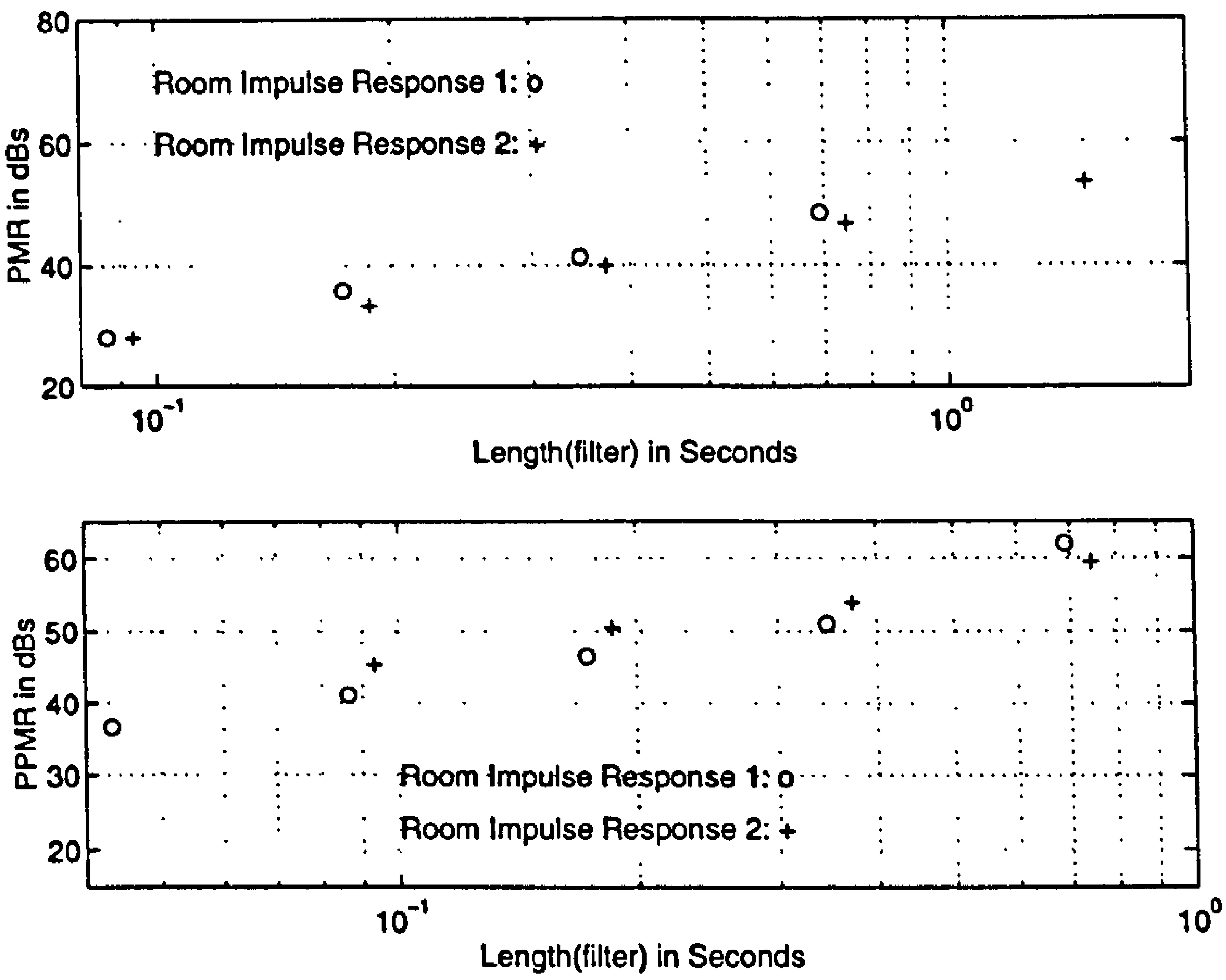
Using this equation the optimum PMR and PPMR for a proposed filter can be examined.

There are are only two parameters that can affect the results of the equation for the optimum filter if the impulse response remains unchanged, either a longer filter can be specified or the system can be "oversampled", ie. the system can operate at a bandwidth greater than that required. There are limits to the application of either technique, perhaps the most significant of these is the processing power available. The cost the proposed system will inherently be linked to the processing requirement.

In the earlier examination (section 7.6.1) it was demonstrated that oversampling the filter has little material effect on performance, thus the only way that a filter can be simply tailored to the quality level required is to adjust its length. (Adding more channels is not within the design brief even though this is probably the most effective performance measure examined.) The performance gain that can be expected with increasing filter length is therefore examined so that an estimate of the required length for sufficiently accurate compensation can be estimated.

As an example consider a real impulse response recorded in the test room and then decimated so that the effective sample frequency is 440Hz. (The impulse measurement was windowed to ensure that the effects of residual noise in the impulse response measurement did not affect the filter generation process.) Filters were developed according to equation 7.6 at lengths ranging from equal to that of the impulse response down to 1/32 of its length. Similar tests were carried out on a number of other real impulse responses over smaller ranges of filter length. On average it appears that the PMR and PPMR reduce by a constant amount for each time the length of the filter is doubled. The results are shown in figure 7.31.

Figure 7.31: The Effect of Filter Length on PMR and PPMR



The extrapolation required to ensure that the PMR and PPMR are greater than -90dB is dangerously large, but essential due to the processing power required to calculate the exact filters. It appears that of the order of 20,000 weights (for 200Hz sample rate) will be required - that is equivalent to 100 seconds. It should be noted that these figures are only to be taken as a rough guide since it is clear that the form of the filter will be determined by the room transfer function, which cannot be predicted usefully.

Very roughly the "pre-filter" algorithm requires the following Multiply-Adds (MAD)s:

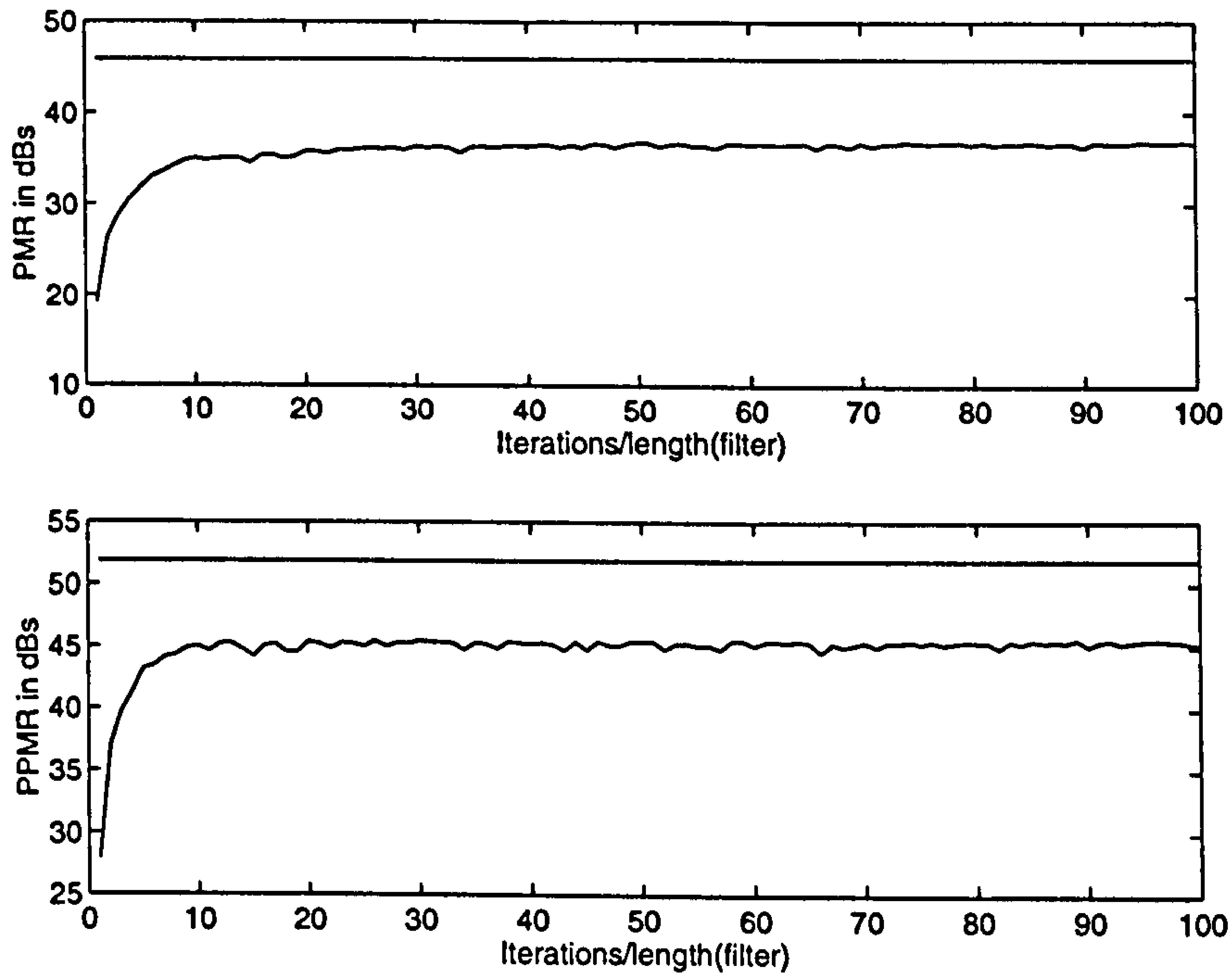
Convolution iteration	length(filter)
Error Calculation	neg
Tap Update	length(filter) * 2

That is roughly 3 * length(filter) MADs are required every sample.

To put this in perspective the system requirement is approximately 12,000,000 MADS/second for a 200Hz sample rate. Given that processor boards are available that can yield 300Mflops [2], say 150MegaMADS, then this processing requirement should not present a problem.

There is an additional problem, the NLMS adaptive algorithm will not reach the exact minimum error solution in a reasonable length of time. While initially the algorithm demonstrates very good convergence performance, the rate reduces rapidly as the solution approaches the optimum. Alternative algorithms or adaptions to the NLMS algorithm that can raise the performance of the adaptive process were discussed in section 7.6. In practice, no algorithm will yield precisely the minimum error solution. Figure 7.32 shows PMR and PPMR for a typical filter adaption, note that only short filters are used for convenience in this demonstration.

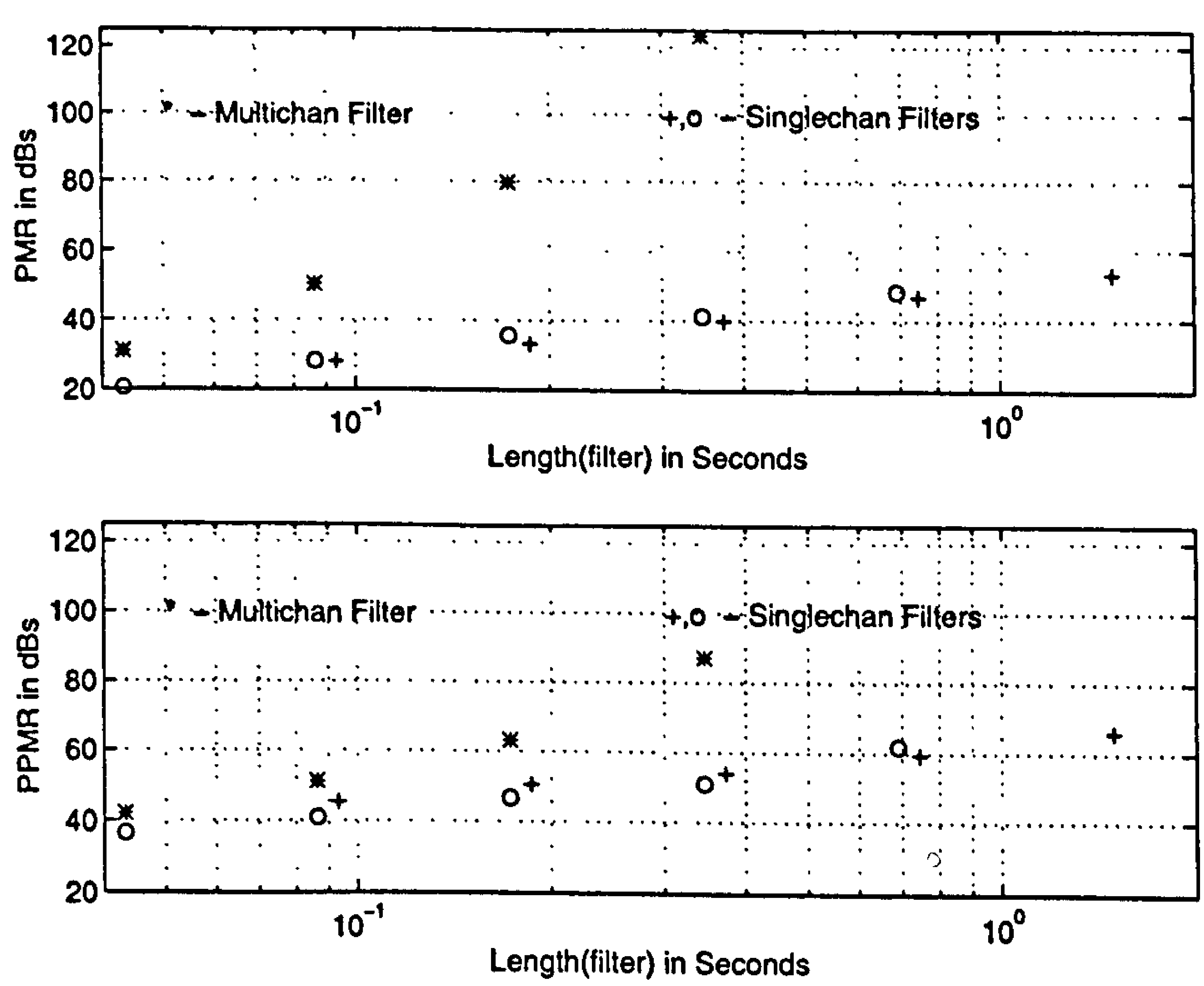
Figure 7.32: PMR and PPMR with Iteration Number



Given that HiFi accuracy is to be attained the extrapolation indicates that very long filters will be required - of the order of 20,000 samples at 200Hz. There are a number of problems that need to be considered with these very long filters. If the delay is set to half the filter length then there will be a 50sec delay between the arrival of the signal and the output. It will also be necessary to store 50 seconds of the signal at full bandwidth so that the bypassed signal components arrive at the correct time. When it is considered that 100 seconds of signal is 4,410,000 samples or 10Mbytes for each channel then at least the cost implications cease to appear as significant. (Current memory prices have RAM at about £25 for 4 Mbytes.) Typically NLMS filters take between 10 and 100 length(filter) iterations to mature, in this case that amounts to between 1000 and 10,000 seconds - over three and a half hours! The filter will have to adapt much faster then this if the results are to be useful.

Since the lengths required are not practical then it will be necessary to make use of the two channels of the stereo system independently. As has been described previously, the two reproduction channels may be used independently so that best use can be made of the "easy" bits of each transfer function to invert. (Considering the series expansion inversion method for a moment it was clear that room modes that resulted in roots close to unity would require long filters to achieve accurate results. These roots are "difficult" to invert. If the second path is available then the same room mode may result in a root further from unity this time and thus offer an "easier" alternative.) This method has been demonstrated to provide a substantial improvement in the PMR and PPMR for a certain length of filter. If this technique is applied then the PMR and PPMR for different filter lengths is given in figure 7.33.

Figure 7.33: PMR and PPMR for Single and Multiple Channels



Now, when the PMR and PPMR are extrapolated towards 90dB the filter length is far more reasonable, less than 1 second would be more than adequate in both cases. In practice a filter length of a few seconds would probably be chosen so that even though the real word performance falls short of the optimum solution the PMR and PPMR remain at better than 90dB.

Overall it appears that a filter of sufficient accuracy could be constructed using NLMS techniques provided that some speed-up method could be devised. To re-iterate, simply achieving PMR and PPMR figure of better than 90dB will not necessarily be sufficient to be acceptable to the HiFi fraternity.

7.7.7 Background Noise

The system uses a microphone to obtain a signal which it compares with the target and produces a filter to make the observed signal as close as possible. If the signal at the sensor is corrupted by noise then the filter will be changed in a spurious attempt to match the signals. Noise will therefore limit the accuracy of the final filter.

Low frequency noise can be a particular problem, for instance closing doors produce a very high SPL at very low frequencies. It might be possible to reduce the scale of this type of problem by limiting the LF signal content.

In practice the limitations will depend on the filter scheme used, the "filtered-x" concept after Widrow is claimed to be more robust in the presence of noise than a pre-filtered method. As previously discussed the noise tolerance of the filtered-x algorithm is indeed better, but even at levels of noise far in excess of those expected in a real HiFi situation the pre-filtered algorithm performed better for realistic numbers of iterations. There are a number of "intelligent" schemes that can reduce the effect of short lived noise, which can be particularly damaging to mature filters. (For instance, if the error value suddenly becomes very much larger adaption can be stopped until either the value drops again or the control logic must conclude that the system has changed dramatically.) It is expected that a number of these controls would be implemented on a commercial system.

7.7.8 Noise Pollution

The inverse filter acts in destructive interference mode, rather than actually removing the reflected sound energy. This means that as the cancellation continues extra energy is added to the system to cancel the effect of the reflections at the listening point - and the effect of the reflections of the previous cancellation signal etc. There will therefore be higher levels of acoustic energy in the room for a given listening level than would have occurred without the inverse filter in operation which will inherently exacerbate any existing problem with noise pollution.

7.7.9 Conclusion

It appears that the most severe limitation to the system is the inherent physical separation between the sensor and the listening point. In practice there is no easy solution to this problem. Provided that the sensor can be close to the listening point then it is possible to consider a system that will operate on a reduced bandwidth, it appears that a realistic limit is of the order of 100Hz.

A possible partial solution to the location of the sensor problem is to make use of two sensors and interpolate towards the listening point(s). (There are two effective listening points to be considered if the sensor/listening point distance is to be reduced significantly below 0.2m.) Provided that the sensors could be positioned on the same axis as the listener's ears and the distances between then various points predicted accurately then it seems likely that some useful interpolation could take place. For instance, considering an acoustic plane wave that is perpendicular to the axis then it is clear that there will be no difference between the output of the sensors and the signal perceived by the listener. If the displacement were in the same direction of the wave then the change would be rapid, but it would still be possible to approximate the signal perceived by the listener with appropriate interpolation. The system would fail to be useful when the wavelength approached half the distance between the two points (Similar to the Nyquist limit).

Unfortunately it is difficult to imagine an acceptable commercial system that could locate the sensors on the same axis as the listener's ears with sufficient accuracy to be useful.

It is generally accepted that it is difficult to achieve an acceptable bass response in a small room [54] and thus this technology may prove beneficial in practice even though it is likely that there will never be a solution to the problem of the distance between the sensor and the listener's ears.

7.8 Summary and Proposal for a Practical System

Recent developments in DSP hardware and signal processing in general mean that it is now realistic to consider systems to compensate for the effects on audio reproduction of acoustic reflections in the listening room. Fixed systems are available now commercially, for instance the device manufactured by SigTech, that are claimed to offer a real improvement in the quality of reproduction. Such fixed systems are not entirely suitable for the consumer market for practical reasons. Adaptive filter technology has, however, progressed to a stage where it is reasonable to consider its application to this problem. Nelson *et al* [44] and others have demonstrated the basic concepts and developed a multi-channel NLMS algorithm for this task but the bandwidth and filter quality of the system were such that no useful appraisal of the results could be obtained.

Consideration of the practical difficulties that will be faced if the system were implemented has indicated that it is unlikely to be a realistic proposition except at the extreme LF end of the audio band. By far the most intractable problem is the distance that must exist between the the sensor(s) and the listener's ears. Even if there is only one listener and the sensor is positioned as close as is practical to the head then the usable bandwidth is about 100Hz. Exceeding this bandwidth or movement of the user will cause the quality of the response to fall as the distance increases. Significant movement of the listener will yield a response that is definitely inferior to the uncorrected response.

Even at frequencies of less than 100Hz there are strong acoustic effects in a normal sized listening room. These low frequency modes are very difficult to deal with using conventional methods (damping the reflections of the walls etc) since the absorption of most realistic surfaces is limited at low frequency. Thus, at low frequency, there may be a real case for the application of this technology in spite of the limited band over which it can operate.

Given that the bandwidth is fundamentally limited to less that 100Hz a great deal of the remaining problems and limitations are simplified. For instance sensor selection becomes much easier as accuracy is only required over one decade and only one sensor is required since the human hearing system is certainly incapable of locating a sound using the clues supplied by the stereo system at frequencies of this order.

It has been demonstrated that only a single channel of filter is required to produce an exact inverse of the room acoustic at a single location provided that the filter is allowed to extend both forward and backward in time to infinity. This result is a direct contradiction of statements made by both Nelson *et al* [44] [46] [45] and Miyoshi *et al* [39]. It has also been demonstrated that there are significant advantages to making use of the two stereo channels independently since this allows much more practical length filters.

Various other methods for improving the performance of the inverse filter have been considered, drawn from the efforts that have been made when working with a telephony problem that is somewhat similar. The performance improvements yielded probably do not warrant their inclusion in any final system, especially since the processing power is not a limiting factor.

There is little doubt that there exists a small market for very high performance HiFi equipment which will pay substantial sums for anything that will improve the quality of reproduction. It is not certain how well the limited bandwidth version that is possible could be marketed even though this is the limit of practical application of this technology.

The bandwidth limitation is due to a physical constraint, it seems exceeding unlikely that there will ever be any solutions developed that can resolve the inherent problem. There is sufficient technology available to build the hardware required for the limited version available now and thus there is little impetus to research techniques for reducing the processing power requirement.

The demonstration that it is possible to produce an exact, if not finite length, inverse to an non-minimum phase system using NLMS techniques directly (provided that a delay is acceptable) may have application in other fields. There seems to be considerable confusion over this matter. There are many compensation problems that have a similar "inverse filter" form to the HiFi acoustic problem discussed here, such as compensators for any systems involving a dispersive channel such as radio transmission etc.

Chapter 8

Conclusions

Investigations have been carried out into three promising applications of DSP in the HiFi/sound reproduction environment. The applications were:

- Using the back EMF to allow feedback control of loudspeakers
- Compensation for Doppler distortion
- Room acoustic compensation

8.1 Using the back EMF to allow feedback control of loudspeakers

The concept of using the back EMF to control the cone velocity of a Loudspeaker has a number of advantages that make it particularly attractive to a commercial application. There is no need to alter current designs in order to make use of the technology and no additional sensors are required external to the normal speaker system. The basic idea is not new, there have been a number of attempts made to make use of the feature. Previous attempts have floundered on the presence of a series resonance, which means that the measured back EMF is not directly proportional to the velocity as might have been hoped.

Any system that hoped to use the back EMF thus has to have some mechanism to deal with this series resonance. It is not possible to set a model for the series resonance in advance since this feature is very sharp and is the result of the cone mass oscillating with the series inductance (of the voice coil) the former cannot be assumed to be stable. The bridge used to obtain the back EMF will also be affected by changes in the coil resistance, which can be expected to be significant.

By applying system identification techniques to mimic the output of one leg of resistive bridge it is possible to get a very good estimate of the speaker bulk parameter model's coefficients. Using this information a bridge can be formed to cancel the effects of both the coil resistance and inductance. Thus the output of the bridge is the back EMF.

Simulations of this technique produced very promising results.

8.2 Compensation for Doppler distortion

Perhaps surprisingly, since the effect is non-linear, it is possible to compensate for the distortion caused by the cone velocity. If there are two signal components to be reproduced on a single driver both will cause the cone to move. If one is assumed to be at a low frequency then due to the mechanics of the operation of a loudspeaker this will cause relatively large cone velocities. The second component at a higher frequency will be Doppler distorted since it is reproduced on a moving surface.

The human hearing system is theoretically only just sensitive enough to detect this distortion so even a relatively small improvement should be enough to render it inaudible.

The compensation scheme is not easily modelled with normal techniques but since the nature of the distortion is quite easy to comprehend it is still possible to design a correction filter. The system requires knowledge of the current cone velocity, this can either be obtained directly or a model can be used. The accuracy of the model is not critical.

Measurements indicated that the application of the filter developed produced an improvement in the level of this type of distortion. Unfortunately it was not possible to produce a convincing subjective demonstration of this technology but this failure was largely attributed to the very basic equipment available for the demonstration.

8.3 Room Acoustic Compensation

When a stereo music signal is reproduced in a room the listener will hear both the direct signal from the loudspeaker and reflections of these signals from every surface in the room. By applying adaptive filter techniques it initially appeared that it might possible to completely cancel the effects of all of these reflections over the full audio bandwidth. After detailed consideration the conclusion that must be drawn is that it is fundamentally not possible to attain useful performance over the full audio bandwidth when practical limitations are considered. The largest stumbling block is the distance that is inherent between the sensing microphone (required for the system to adapt) and the listening point. There are additional problems when the stereo content of the signal is considered.

Previous researchers in this area did not apparently consider these limitations, which would not have been apparent due to the very limited bandwidth and the unrealistically short filters that they used to test the concept. The problem is not inherent when fixed filters are used, such as the SigTech device, since the sensing microphones can in theory occupy the listening position. Even though there is no fundamental limitation SigTech found it necessary to reduce the bandwidth over which full correction was applied for a practical system to operate satisfactorally.

The constraints indicate that it is possible to use the techniques up to a maximum frequency of the order of 200Hz, the exact limit depending on the layout of the target system. At present commercial application of the technique is unlikely since the limitations are serious. There are, however, potential gains to applying this system since even below 200Hz the effects of the listening room are likely to be significant and damping reflections by mechanical means (insulation on the walls etc) becomes much less effective at low frequencies.

There appears little future for further work in this area since the problems faced are not likely to become tractable even with new techniques.

Part of the initial brief was to consider the application of this technology to live music situations for sound reinforcement. Here the problems faced tend to be far more serious and it is highly unlikely that any application of this technology will be possible at all.

Bibliography

[1] *The International Dictionary of Physics and Electronics.* Van Nostrand.

[2] Advertisements. Signal processing magazine, October 1993.

[3] R. Allison and E. Villchur. On the magnitude and audibility of fm distortion in loudspeakers. *Journal of the Audio Engineering Soc.*, 30(10).

[4] D.W. Armitage. Servo-control of a moving coil loudspeaker. Hons. dissertation, UCNW, Bangor, 1972.

[5] M.R. Asharif and F. Amando. Frequency bin adaptive filtering (fbaf) algorithm and its application to acoustic echo cancelling. *IEICE Transactions on Communications Electronics Information and Systems*, 74(8), 1991.

[6] D. Birt. Loudspeaker power amplifiers with load-adaptive source impedence. *Journal of the Audio Engineering Society*, 36(7/8), 1988.

[7] J. Borish and J.B. Angell. An efficient algorithm for measuring the impulse response using pseudorandom noise. *Journal of the Audio Engineering Society*, 31(7):478 – 488, 1983.

[8] J.M.P. Borrallo and M.G. Otero. On the implementation of a partitioned block frequency-domain adaptive filter (pbfdaf) for long acoustic echo cancellation. *Signal Processing*, 27(3), 1992.

[9] Bronshtein and Semendyayev. *Handbook of Mathematics.* Van Nostrand, 1979.

[10] P. Lin Chang and B. C.G., Yeh. Inverse filtering of a loudspeaker and room acoustics using time-delay neural networks. *Journal of the Acoustic Society of America*, 95(6):3400 – 3408, 1994.

[11] P. M. Clarkson, J. Mourjoopulos, and J. K. Hammond. Spectral, phase, and transient equalisation for audio systems. *Journal of the Audio Engineering Soc.*, 33:127–132, 1985.

[12] Martin Colloms. *High Performance Loudspeakers.* Pentech, 1985.

[13] AJ Compton. *Basic Electromagnetism.* Van Nostrand Reinhold.

[14] a. Davies. The room environment. *Collection, title lost.*

[15] Richard C. Dorf. *Modern Control systems.* Addison Wesley, 1989.

[16] John Downer. *Supersense.* BBC Books, 1988.

[17] C. Dunn and M.O.J Hawksford. Distortion immunity of mls-derived impulse response measurements. *Proceedings of the Institute of Acoustics*, 13(7):191–206, 1991.

[18] Eminence. *High Power Loudspeaker Enclosure Design and Construction.* Adam Hall Supplies Ltd, 1991.

[19] K. D. Farnsworth, P. A. Nelson, and S. J. Elliott. Equalization of room acoustic responses over spatially distributed regions. *Proceedings of the Institute of Acoustics, Autumn Conference*, 1985.

[20] W.A. Frank. An efficient approximation to the quadratic Volterra filter and its application in real-time loudspeaker linearization. *Signal Processing*, 45:97–113, 1995.

[21] K. Fujii and J. Ohga. A fast algorithm suitable for acoustic echo canceller. *IEICE Transactions on Fundamentals of Electronics Communications and Computer Sciences*, VE75A(11), 1992.

[22] Stanley A. Gelfand. *Hearing, An Introduction to psychological and Physiological Acoustics.* Marcel Dekker, Inc., 1990.

[23] R. Genereux. Adaptive filters for loudspeakers and rooms. *Audio Engineering Society UK DSP Conference*, pages 175–195, 14-15 September 1992.

[24] R. Genereux. Adaptive fileters for loudspeakers and rooms. *93rd Convention of the Audio Engineering Society*, 1992.

[25] A. Gilloire and M. Vetterl. Adaptive filtering in subbands with critical sampling - analysis, experiments, and application to acoustic echo cancelling. *IEEE Transactions on Signal Processing*, 40(8), 1992.

[26] S. Gudvangen and S.J. Flockton. Comparison of pole-zero and all-zero modelling of acoustic transfer functions. *Electronics Letters*, 28(21), 1992.

[27] D. Guicking, K. Karcher, and M. Rollwage. Coherent active methods for applications in room acoustics. *Journal of the Acoustic Society of America*, 78(4), 1985.

[28] E. Hansler. The hands free telephone problem - an annotated bibliography. *Signal Processing*, 27(3), 1992.

[29] F.R. Heath. Loudspeaker feedback systems. Hons. dissertation, UCNW, Bangor, 1970.

[30] C. Herley. *Wavelets and Filter Banks.* PhD thesis, Columbia University, New York, NY, April 1993.

[31] J.E. Kimberly. Development of servo-control of a moving coil loudspeaker. Hons. dissertation, UCNW, Bangor, 1973.

[32] Fray Kinsler. *Fundamentals of Acoustics.* Wiley, 1962.

[33] R.H. Kwang. A variable step size lms algorithm. *IEEE Transactions on Signal Processing*, 40(7), 1992.

[34] Lennart Ljung. *System Identification Theory for the User.* Prentice Hall, 1987.

[35] Jeff Maggard. Dat-heads frequently asked questions - microphone edition. Public domain information, Mailing List (DAT-Heads-Request@fuggles.acc.Virginia.EDU), 1993.

[36] S. Makino and Y. Kaneda. Exponentially weighted step-size projection algorithm for acoustic echo cancellers. *IEICE Transactions on Fundamentals of Electronics Communications and Computer Sciences*, VE75A(11), 1992.

[37] S. Makino, Y. Kaneda, and N. Koizumi. Exponentially weighted stepsize nlms adaptive filter based on the statistics of a room impulse response. *IEEE Transactions on Speech and Audio Procession*, 1(1), 1993.

[38] D.Y. Middleton. Loudspeaker motional feedback. Hons. dissertation, UCNW, Bangor, 1966.

[39] Masoto Miyoshi and Yutaka Kaneda. Inverse filtering of room acoustics. *IEEE Trans. Acoustics, Speech and Signal Processing*, 36(2), 1988.

[40] P. M. Morse and K.U. Ingard. *Theoretical Acoustics.* Princeton University Press, 1986.

[41] A.S. Munro and P.J. Pyatt. Principles and practice in maximum length sequence analysis. *Proceedings of the Institute of Acoustics*, 13(7):57–64, 1991.

[42] H Negishi. Wide imaging stereo (wis) technology. *Electronics and Communication Engineering Journal*, 5(1), 1993.

[43] P. A. Nelson, S. J. Elliott, and I. M. Stothers. Improvements in or relating to sound reproduction systems. The Patent Office, January 1990. International Publication Number is: WO 90/00851.

[44] P.A. Nelson, S.J. Elliott, and H. Hamanda. Adaptive inverse filters for stereophonic sound reproduction. *IEEE Transaction on Signal Processing*, 40(7), 1992.

[45] P.A. Nelson, S.J. Elliott, and H. Hamanda. Inverse filters for multi-channel sound reproduction. *IEICE Transaction Fundamentals*, E75-A(11), 1992.

[46] P.A. Nelson, F. Orduna-Bustamante, and H. Hamanda. Inverse filter design and equalisation zones in multichannel sound reproduction. *IEEE Transactions on Speech and Audio Procession*, 3(3), 1995.

[47] F. Ordunabustamente and P.A. Nelson. An adaptive controller for the active absorption of sound. *Journal of the Acoustic Society of America*, 91(5), 1992.

[48] J.J. O'Reilly. *Telecommunication Principles.* Chapman and Hall, 1984.

[49] T. Petillion, A. Gilloire, and S. Theodoridis. A fast newton transversal filter: An efficient scheme for acoustic echo cancellation inmibile radio. *IEEE Transactions on Signal Processing*, 42(3), 1994.

[50] John R. Pierce. *The Science of Musical Sound.* W.H.Freeman, 1992.

[51] D.D Rife. Transfer-function measurement with maximum length sequences. *Journal of the Audio Engineering Society*, 40(4):260–276, 1992.

[52] Elain Saxton. Digital versatile disc. *New Electronics On Campus*, Summer 1996, 1996.

[53] H. Schutze and Z. Ren. Numerical characteristics of fast recursive least-squares transversal adaptive algorithms - a comparative study. *Signal Processing*, 27(3), 1992.

[54] Ian R. Sinclair. *Audio Electronics Reference Book.* BSP Professional Books, 1989.

[55] R.H. Small. Direct radiator loudspeaker system analysis. *Journal of the Audio Engineering Soc.*, 20(5), 1972.

[56] J.W. Strutt and Baron Raleigh. *The Theory of Sound*, volume 2. Macmillan co, London.

[57] Sigtech (Cambridge Signal Technologies). http://www.sofgry.com/sigtech/index.html. Www page.

[58] A.N. Theil. Loudspeakers in vented boxes: Part ii. *Journal of the Audio Engineering Soc.*, 19:471–483, 1971, June.

[59] M. Tohyama, R. H. Lyon, and T. Koike. Pulse waveform recovery in a reverberant condition. *Journal of the Acoustic Society of America*, 91(5), 1992.

[60] Klippel W. The mirror filter - a new basis for reducing nonlinear distortion and equalizing response in woofer systems. *Journal of the Audio Engineering Soc.*, 40(9).

[61] J White. Stereophonic sound in the home - hearing is believing. *Proceedings of the Institute of Acoustics*, 13(7):339 – 351, 1991.

[62] Bernard Widrow and Samuel D. Stearns. *Adaptive Signal Processing.* Prentice Hall, 1985.

[63] William W.Seto. *Acoustics.* Schaum, 1971.

[64] H. Yasukawa and S. Shimanda. An acoustic echo canceller using subband sampling and decorrelation methods. *IEEE Transactions on Signal Processing*, 41(2), 1993.

[65] H. Yasukawa and S. Shimanda. An acoustic echo canceller using subband sampling and decorrelation methods. *IEEE Transactions on Signal Processing*, 41(2), 1993.

[66] B. Zoltogorski. Moving boundary-conditions and nonlinear propagation as sources of nonlinear distortions in loudspeakers. *Journal of the Audio Engineering Soc.*, 41(9).

[67] D. T. N. Williamson More views on loudspeaker damping. *Wireless World*, 53(10), 1947.

[68] R. B. McGreggor Letter - improving loudspeaker response with motional feedback. *Electronics (McGraw-Hill)*, 24(5), 1951.

[69] R. L. Tanner Improving loudspeaker response with motional feedback. *Electronics (McGraw-Hill)*, 24(3), 1951.

Chapter 9

Appendices

9.1 Basic Methods used for Experiments

In this section the methods used to measure the room's impulse response and generate the inverse filter are considered in detail. Listings of the Matlab code used may be found in the appendix, section9.4. This section deals with the practical aspects of the method rather than the mathematical justification, it is intended to make reproducing the results less difficult.

The basic method employed has been to measure the impulse response in question by making use of the PRBS and correlation techniques and then use this information to simulate the effects of the room on a signal. The advantage of this method is that it allows all processing to be carried out off-line, reducing the processing power requirement and allowing more accurate monitoring of the processes. (It has also allowed processing to be carried out on a general purpose computer with the mathematics package "Matlab" rather than requiring all code to be written at assembly level for the DSP) The disadvantage of this method is that the room impulse response model will be entirely linear, thus the effects of the non-linearities that may be present cannot been seen in the simulations. In practice, as has been noted before, the room acoustic effects are expected to be strongly linear and thus this limitation is not likely to be significant.

9.1.1 Measuring the Room's Impulse Response

Of the methods available for measuring the response of a room the one that appeared to yield the best balance of accuracy, noise tolerance and absolute phase information was the Pseudo Random Binary Sequence (PRBS) method. Initial attempts to make use of a gain/phase analyser floundered on the instruments inability to measure the absolute phase of the system. (Attempts were made to unwrap the relative phase information provided by the unit and thus guess the most likely absolute phase value. The results yielded became poor at quite modest frequencies. There were additional problems with noise pollution, speed and frequency resolution. The PRBS method was not chosen initially since the required equipment was not available.)

The PRBS method is well documented and has been used for this type of application for a number of years [7], [51], [17] and [41]. This method uses the concept that it is possible to produce a binary sequence that simulates white noise up to a certain limit frequency and has an autocorrelation function that is an impulse in time. If such a signal is passed through a perfectly linear and noise-free system and the result is circularly cross correlated with the original then the output is exactly identical to the Periodic Impulse Response (PIR) of the system.

There are a number of very significant advantages to using the PRBS method over the more direct PIR technique, the most important of which is that the power in a PRBS signal is vastly greater than that in a PIR train. This translates to good noise immunity and/or the ability to operate the system at lower levels and thus reduce the possibility of distortion. Other features of this method are:

- DC Offset

  The PRBS method severely attenuates any DC offset signal that is present.

- Truncation (Time Aliasing) Effects

  The PRBS method produces a periodic result. Provided that the measured response is shorter than the length of the sequence then there is no truncation effect and thus no data window is needed [41]. If the response measured is longer than the sequence the tail will be added to the start. In practice the sequence length is chosen to be as long as practical to reduce both of these effects.

- Transient Noise Immunity

  Transient noise is transformed into low level benign noise spread across the response thus the immunity is good. (Similar transient noise applied to the PIR method would be recorded directly in the output and the effects could only be reduced by averaging.

- Non-Linearity in PRBS measurements

  The effect of non-linearities is to cause quasi-stationary fixed pattern noise to appear in the response [41]. If this is a problem the results can be reduced by using longer sequences than required so that the tail can be discarded and with this most of this type of noise. (Ideally the effects of non-linearities should be modelled correctly in the impulse response rather than discarded as noise. In practice this would mean an impulse response that changed with input level, a complication not justified by largely linear systems such as room acoustics yields.)

Even using the PRBS technique background noise can remain a problem, particularly low-frequency noise from air conditioning, traffic, doors etc. (Low-frequency noise is a particular problem since it is less affected by intervening structures.) There are two methods that can be use to reduce the effects of this type of noise:

- Averaging

  It is possible to average a number of results to increase the effective power of the desired signal. (The response is deterministic and will accumulate whereas the noise component is random and hence will tend to cancel.)

- Pre-emphasis

  PRBS sequences have energy characteristics similar to white noise which means that for wide-band measurements there is little energy at low frequencies. This imbalance can be addressed by pre-filtering the PRBS sequence provided that the reverse procedure is carried out with the recorded signal.

Details of the System Used

The impulse response measurement software was written especially for this research. It was written in the form of Matlab functions and was designed to work with the digital audio interface provided by Berkeley Camera Engineering (Berkeley, USA). Some alterations were required to the DSP code and access functions provided by Berkeley Camera in order to allow the best use to be made of the ADC and DAC. (The ADC and DAC were both 18 bit devices, the DAC was provided by Berkeley Camera and the ADC sourced from Crystal Semiconductor Corporation with support systems built especially for this project.) A full listing of the code is provided in appendix, section 9.4.

The process starts by building the signal to be transmitted. This needs to consist of at least one sequence to "initialize" all of the signal in transit. Due to the non-minimum phase nature of the room impulse response it was chosen to use a pair of sequences to form this pre-amble. Following this there needs to be at least one pair of sequences for the actual measurement. In practice a number of repeats can be selected to allow averaging of the response obtained in order to reduce the effects of noise. No special considerations are required at the end of the sequences.

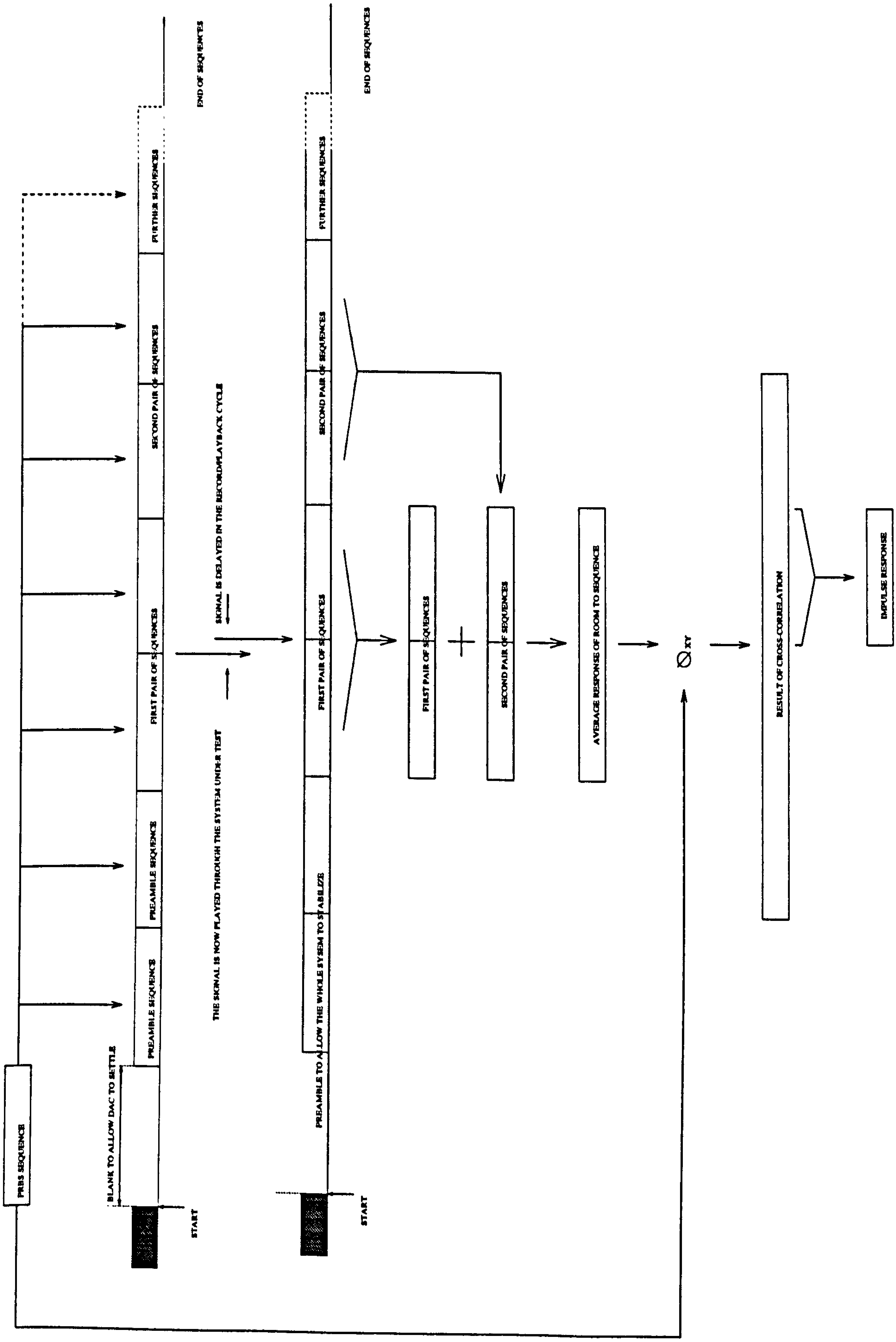
The signal is then passed to the DAC and the result is recorded. The first stage of the processing is to discard the results of the pre-amble; or more accurately the samples recorded at the same time that the pre-amble was being transmitted since the end of the pre-amble will not have reached the measuring microphone until significantly after it was transmitted. (This is the reason for using two sequences as a pre-amble instead of the minimum of one required.) Next the recorded signal is split into lengths according to the number of pairs of sequences transmitted. The lengths of signal are added, to reduce the effects of noise, and then the result is cross-correlated with the original PRBS sequence. Note that the averaging process here occurs before the cross-correlation; since the correlation process requires a very considerable amount of processing time at the sequence lengths required for room acoustics. The advantages, if any, of averaging after cross-correlation, as it is understood that [41] suggests, are not clear.

An additional complication is required if averaging over a large number of sequences due to the limited memory of the machine. It is necessary to construct shorter sequences, pass them through the system, split and average the results and then repeat this process until the necessary number of iterations have been attained. These have been termed cycles. For the maximum efficiency the number of repeats (iterations in every cycle) should be maximized for the available memory and the number of cycles set then to give the desired total number of iterations. This process is shown in figure 9.1.

The PRBS sequence must take longer to reproduce than it takes for the impulse response of the system to fall below the threshold of interest; if it does not then the tail of the impulse response will be folded back onto the body. For the room impulse response the signal roughly decays exponentially so a certain amount of this "aliasing" type effect is inevitable. In practice, with the room used for experiments, the period required is about 300ms. With a sampling rate of 48kHz this represents about 15000 samples.

The noise present in the recording system and the background acoustic noise present in the room represent too much distortion of the impulse response to be ignored. This means that a large number of iterations is required in order to attain reasonable performance. Pre-emphasis was not used in the system although in retrospect the potential performance gain would have probably warranted the effort. A great deal of the noise appeared to originate from the microphone pre-amp which is of an old design and is not really suitable for this type of measurement.

Figure 9.1: Obtaining the Room Impulse Response



The hardware used for the record and playback also introduces a fixed delay into the signal which is of no interest, this is compensated for in the measurement program.

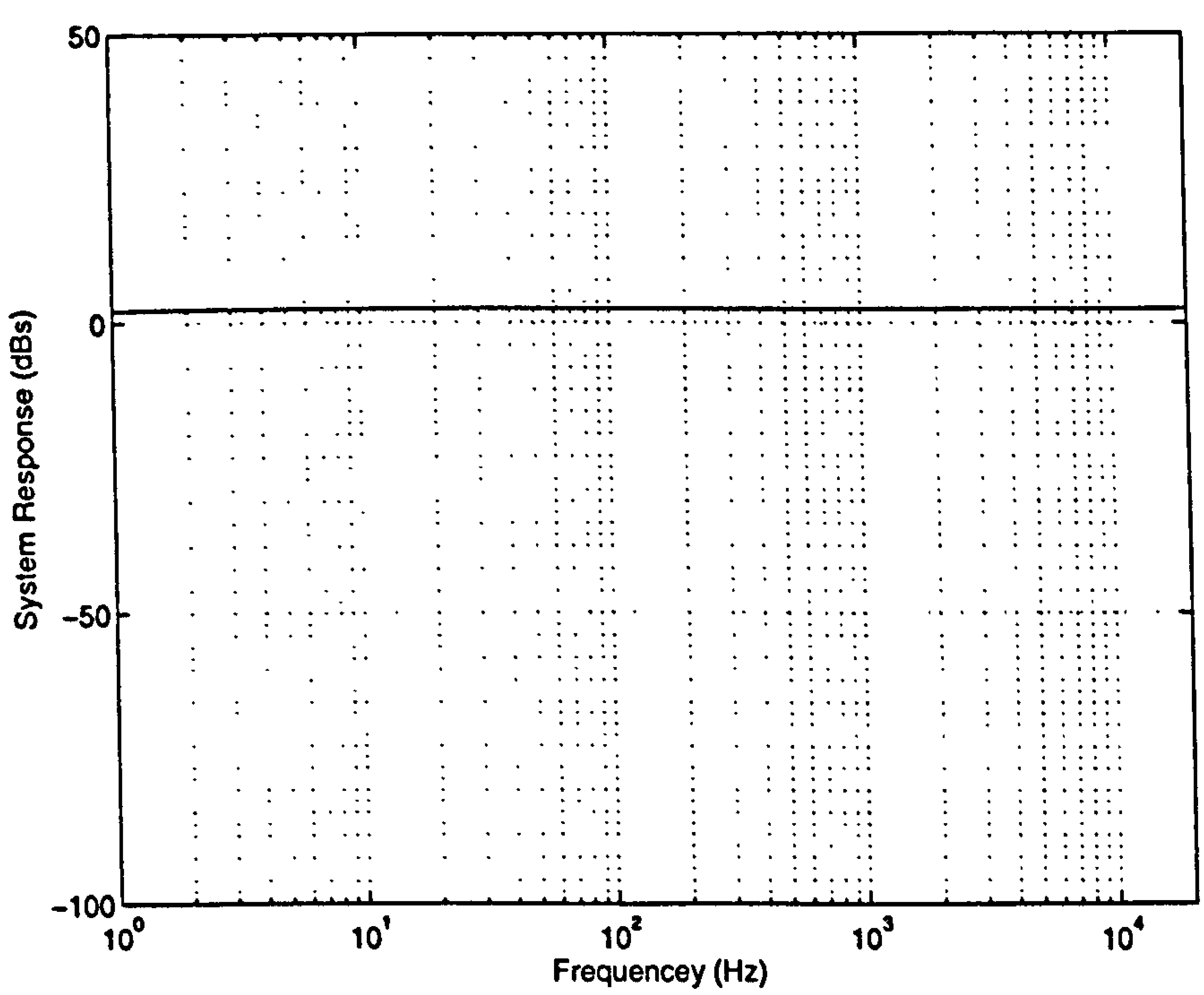
Tests of the impulse measurement system using an LCR network with a predictable impulse response confirmed that the practical system performed as the theory suggested it should.

It is possible to alter the technique to measure two or more responses simultaneously provided that different PRBS sequences are used. (For a given sequence length there are normally a number of different sequences possible and each of these is orthogonal to every other. Cross-correlating two orthogonal sequences will produce no output and thus the sequences can be used simultaneously and the results will remain completely independent.)

Another option is that the signal after one process may be cross correlated with the result of another to yield the relative impulse response.

The system was tested by feeding the output of the DAC into the ADC direct. The performance obtained was very satisfactory. (Note that the scale was chosen to be consistent with the remainder of plots of this type in this document, to allow for easy comparison.) The results can be seen in figure 9.2.

Figure 9.2: System Response of DAC and ADC only



The effects of noise were observed by simply recording the ambient noise and the cross-correlating this signal with a suitable PRBS sequence. Despite the averaging techniques used the noise floor, particularly at LF, was of the order of -70dB although this varied with the highly irregular noise. This is short of the -90/-100dB figure that is normally associated with HiFi equipment but it proved adequate for the research. Pre-emphasis could have been tried but that would not have helped with one of the strongest sources of noise, the microphone pre-amp, which appeared relatively white.

9.1.2 Calculating the Filter

Introduction to the Problem

Consider the acoustic system represented in figure 9.3, consisting of a filter followed by a loudspeaker and microphone in a room. The impulse response of the acoustic part of the signal path is $g_{(t)}$, which corresponds to a transfer function of $G_{(\omega)}$. It would at first appear that the ideal system can be constructed by setting a filter $h_{(t)}$ such that the response of the system is given by:

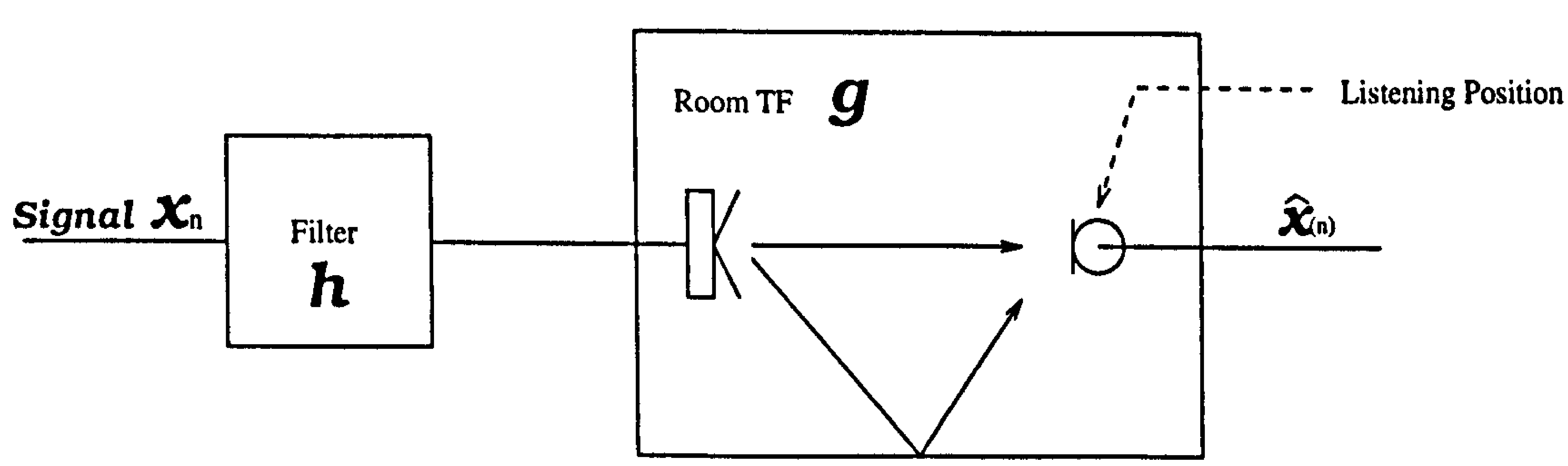
$$\delta_{(t)} = h_{(t)} * g_{(t)} \tag{9.1}$$

Which can be expressed in the frequency domain as:

$$1 = H_{(\omega)}G_{(\omega)} \tag{9.2}$$

In practice this is not a stable process due to the nature of $g_{(t)}$. Producing $\delta_{(t)}$ from the convolution requires that the value of $g_{(t)}$ at $t = 0$ is non-zero. Since there is inherently a transit time for the sound to travel between the loudspeaker and the microphone this condition cannot be satisfied.

Figure 9.3: The System



This problem can be expressed in discrete time in terms of z^{-1}, where z^{-1} is a unit delay. Given that the room impulse response can be expressed as a polynomial in z^{-1}, to the limits of a sampled data system:

$$H_{(z^{-1})} = 1/G_{(z^{-1})} \tag{9.3}$$

$$G_{(z^{-1})} = g_0 + g_1 z^{-1} + g_2 z^{-2} + \cdots \tag{9.4}$$

$$H_{(z^{-1})} = h_0 + h_1 z^{-1} + h_2 z^{-2} + \cdots \tag{9.5}$$

$$D_{(z^{-1})} = d_0 + d_1 z^{-1} + d_2 z^{-2} + \cdots = 1 \qquad \Rightarrow d_0 = 1 \tag{9.6}$$

$G_{z^{-1}}$ could also be expressed as a fraction:

$$G_{(z^{-1})} = \frac{f1_{(z^{-1})}}{f2_{(z^{-1})}} \tag{9.7}$$

When expressed in this way, the function $f1_{(z^{-1})}$ is unity and $f2_{(z^{-1})}$ is a non-unity function of z^{-1} then $G_{(z^{-1})}$ is an IIR filter. Conversely, if $f2_{(z^{-1})}$ is unity and $f1_{(z^{-1})}$ is non-trivial then $G_{(z^{-1})}$ is an FIR filter. It is normal to express adaptive filters in terms of FIR filters only since this type of filter are much less prone to instability due to problems that are difficult to predict such as computational error growth etc. It is possible to make use of IIR filters in adaptive systems with care and intuitively it seems that this type of filter should be able to yield a better match to the infinite responses that characterize this type of system.

Given that $g_0 = 0$, $g_1 = 0$, etc. up to the transit time, ie. $G_{(z^{-1})}$ is non-minimum phase, h_0 etc. will not be finite. This is obvious since the system impulse response is given by $H_{(z^{-1})} * G_{(z^{-1})}$ and thus d_0 is $h_0 g_0$. For $d_0 = 1$ and $g_0 = 0$ h_0 is not finite. The same result can also easily be demonstrated by long division, which is equivalent to deconvolution in the time domain.

There is a solution to this problem that does not require the complications of Multiple-input/output INverse Theorem (MINT) [39] and should be able to provide an exact inverse provided that the filter length is finite. (Contrary to Nelson *et al*'s statement in [44].) It is surprising that Nelson *et al* stated that the problem was unsolvable since they proceeded to use the solution, a modelling delay, in their proposed system.

Demonstrating how the inclusion of a modelling delay will definitely allow the filter to converge is considered in more detail in section 9.2 in this chapter which deals with the convergence of the coefficients of the inverse filter. If the requirement (ie. typically $\delta_{(t)}$) is delayed (so that it becomes typically $\delta_{(t-\tau)}$) such that the coefficients of $G_{(z^{-1})}$ are no longer zero where the response is required then there are real values of h_0 etc that will produce an output. For convenience this "modelling" delay is written as Δ samples.

$$H_{(z^{-1})} = z^{-\Delta}/G_{(z^{-1})} \tag{9.8}$$

The delay can be longer than the minimum required, in fact it is important that the modelling delay is longer than that required for stability if a good inverse is to be found.

The Matrix Method

The methods used to estimate the inverse filters of the FIR type are all essentially similar. All methods must be inexact since the inversion of even one root of the room transfer function will result in a filter of infinite length. Methods are typically based on the Least Mean Square (LMS) family. The most practical method is the Normalized Least Mean Squares (NLMS) algorithm and its many variants.

The simplest way to deal with the problem is to express it in terms of standard matrix operations. When expressed in terms of matrices it is obvious that the problem is over-defined, and hence there will be no exact solution. Matlab has the ability to deal with over-defined matrix problems, by making use of the least mean squares technique, but all of this is transparent to the user. Thus, if the problem were to be expressed in terms of matrices, then Matlab would provide a convenient and high level method of determining the required inverse filter off-line.

The derivation starts by returning to:

$$\delta_{(t-\tau)} = h_{(t)} * g_{(t)} \tag{9.9}$$

Where $\delta_{(t-\tau)}$ is the desired function, $h_{(t)}$ is the filter and $g_{(t)}$ is the room impulse response.

Given $h_{(nT)}$ represents a discrete time approximation to $h_{(t)}$ etc. this equation can be written:

$$d_{(nT-m)} = h_{(nT)} * g_{(nT)} \quad for \quad d_{(nT-m)} = \left\{ \begin{matrix} 1 & n = m \\ 0 & n \neq m \end{matrix} \right\} \tag{9.10}$$

If these signals are truncated an approximation to the expression above is obtained. Since this approximation contains only finite vectors it is amenable to calculation.

$$\mathbf{d_k} = \mathbf{h_k} * \mathbf{g_k} \tag{9.11}$$

Where

$$\mathbf{d_k} = \begin{bmatrix} d_0 \\ \vdots \\ d_m \\ \vdots \\ d_{L_g+L_h+1} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad \mathbf{g_k} = \begin{bmatrix} g_0 \\ g_1 \\ \vdots \\ g_{L_g} \end{bmatrix} \qquad \mathbf{h_k} = \begin{bmatrix} h_0 \\ h_1 \\ \vdots \\ h_{L_h} \end{bmatrix} \tag{9.12}$$

given

$$h_{(L_h T)} = h_{L_h} \tag{9.13}$$

The convolution can be expressed in matrix form as follows:

$$\mathbf{d_k} = \begin{bmatrix} g_0 & & 0 \\ g_1 & g_0 & \\ \vdots & g_1 & \ddots \\ g_{L_g} & \vdots & g_0 \\ & g_{L_g} & g_1 \\ 0 & & \ddots & \vdots \\ & & g_{L_g} \end{bmatrix} \begin{bmatrix} h_0 \\ h_1 \\ \vdots \\ h_{L_h} \end{bmatrix} \tag{9.14}$$

$$\mathbf{d_k} = \mathbf{G_k h_k} \tag{9.15}$$

where $\mathbf{G_k}$ is a $(L_g + L_h + 1) \times (L_h + 1)$ matrix.

This is an over-determined set of simultaneous linear equations of the form:

$$\mathbf{Ax} \simeq \mathbf{b} \tag{9.16}$$

It is not possible to calculate an exact solution to an equation of this form unless some of the linear equations it contains are replicated. If a least squares method is applied it is possible to arrive at the best approximate solution. Matlab is able to produce such an approximate solution at matrix level and therein lies the value of this representation.

The solution $\mathbf{h_k}$ cannot be calculated by finding $\mathbf{G_k}^{-1}$ since the matrix is not square and thus the inverse is not defined. Further, since for this problem generally the rank of the augmented matrix $[\mathbf{G_k d_k}]$ is greater than the rank of $\mathbf{G_k}$ there can be no unique solution.

Both Matlab and the NAG libraries provide routines that can deal with the over-determined case. The filters calculated by these routines have proved very accurate but the method suffers from a computational cost in terms of time taken and memory required that appears to be proportional to L_h^3, which rapidly becomes prohibitive.

There are many reasons why the least squares process could not be applied to a real system at such a high level, perhaps the most important of which is that this method requires that the room transfer function is known. It is possible to measure the room transfer function either by making use of PRBS methods or simply using the audio signal that is being reproduced at the time, though it will take far longer to achieve comparable results using the latter method. Obviously it would be far more desirable if the filter could be approximated directly. Other problems are the memory requirements and the length of time required to generate an answer. Generating the first solution is less of a problem than this concept's inability to work with the previous answer and update it.

Generally this method is convenient, easy to understand and quick implement. Thus it was used for testing purposes, studying the form of the inverse filter and the evaluating potential accuracy that can be achieved.

Note that by finding the point at which the gradient of the error surface is zero it is possible to produce an exact expression for the optimum filter. Note that this is only the optimum solution, there is no perfect solution to such a problem. Since, where it exists, the inverse of a square matrix is unique the optimum filter must also be unique.

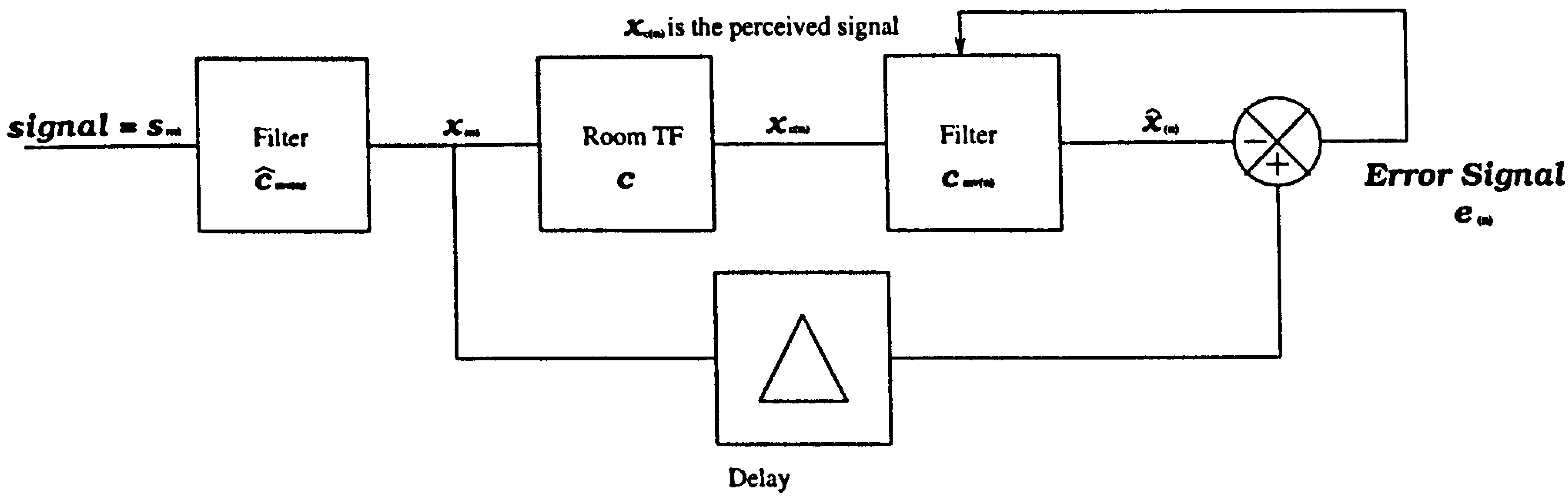
$$h_k = \left[G_k^T G_k\right]^{-1} G_k^T d_k \qquad (9.17)$$

Using the Least Mean Squares Method Directly

The solution to the problems of practically applying the matrix method is to use the Least Mean Squares (LMS) method directly. For this problem the Normalized Least Mean Squares (NLMS) variant has been chosen, the difference being that the adaption step is normalized so that adaption continues at the maximum rate for guaranteed stability even if the signal level falls.

There are two options for the basic architecture, either the room precedes the inverse filter or vice versa. The advantages of each method are considered in depth later in this chapter but at this stage it is sufficient to say that the "pre-filter" method (where the room precedes the inverse filter) was chosen as the bench-mark since this offered much faster adaption rates. Figure 9.4 explains the concept.

Figure 9.4: Proposed Algorithm



The algorithm remains exactly as for the basic NLMS inverse filter problem but with the addition of another inverse filter preceding the whole. The inverse filter is not inside the adaption loop so it will have no effect on the convergent properties of the algorithm.

Initially the filter $\hat{c}_{inv(n)}$ is set to an impulse, so that it has no effect on the signal. The input signal is therefore effectively fed direct to the room (and the listener). The sensor in the room picks up the signal as it arrives at the listening point and passes this to the filter $c_{inv(n)}$, which begins to learn to be an inverse of the room. It learns because its output is compared with a delayed version of the signal and its weights are changed using the NLMS algorithm on the basis of the error. When a monitoring program decides that the filter $c_{inv(n)}$ has become sufficiently mature its weights are copied to $\hat{c}_{inv(n)}$. If desired the filter $c_{inv(n)}$ could be toned down at this stage by adding to the weight value corresponding to the modelling delay.

x = signal vector

$x_{(n)}$ = current signal value

x_c = signal vector filtered by room

$x_{c(n)}$ = current filtered value

$\hat{x}_{(n)}$ = current signal approximation.

c = the room impulse response

$c_{(i)}$ = coefficient in c

$c_{inv(n)}$ = present filter

$c_{inv(n,i)}$ = i^{th} element in the present filter

$c_{inv(n+1)}$ = new filter

s = signal vector

$s_{(n)}$ = current signal

$\hat{c}_{inv}$ = inverse filter used acoustically

$$\begin{bmatrix} \\ c_{inv(n+1)} \\ \\ \end{bmatrix} = \begin{bmatrix} \\ c_{inv(n)} \\ \\ \end{bmatrix} + \frac{\beta e_{(n)}}{K + \sum_{i=0}^{I-1} \left(x_{c(n-i)} \right)^2} \begin{bmatrix} \\ x_c \\ \\ \end{bmatrix} \tag{9.18}$$

And

$$x_{(n)} = \sum_{i=0}^{I-1} s_{(n-i)} \hat{c}_{inv(i)} \tag{9.19}$$

$$x = \begin{bmatrix} \sum_{i=0}^{I-1} x_{(n-i)} \hat{c}_{inv(i)} \\ x_{(n-1)} \\ \vdots \\ \vdots \\ x_{(n-I+1)} \end{bmatrix} \tag{9.20}$$

$$x_{c(n)} = \sum_{i=0}^{I-1} x_{(n-i)} c_{(i)} \tag{9.21}$$

$$x_c = \begin{bmatrix} \sum_{i=0}^{I-1} x_{(n-i)} c_{(i)} \\ x_{c(n-1)} \\ \vdots \\ \vdots \\ x_{c(n-I+1)} \end{bmatrix} \tag{9.22}$$

$$e_{(n)} = x_{(n-delay)} - \hat{x}_{(n)} \tag{9.23}$$

$$= x_{(n-delay)} - \sum_{i=0}^{I-1} x_{c(n-i)} c_{inv(n,i)} \tag{9.24}$$

In practice the signal used was a fragment of a PRBS sequence since this type of sequence is easy to generate, has reasonable spectral qualities and is entirely repeatable. (It is important that the tests are repeatable if useful performance comparisons are to be made.) The use of real signal fragments would have been more appropriate but the storage required and the speed penalty incurred in accessing this sequence would have been severe.

The algorithm was coded for Matlab and in Fortran - since the high power computers used to evaluate full filters quickly could only vectorize Fortran code. Note that there are a number of different versions of the Fortran code, this is due to the highly unpredictable nature of the language implementation. Each platform that the code was required to run on needed slightly different versions of the code.

The version designed to run on the high power computer, a VPS vector machine based at Manchester, was tested and the analysis program reported that the code vectorized very well.

All versions of the code were tested against each other to ensure that they produced consistent results.

9.2 Convergence of Coefficients of an Inverse Filter

In this section the requirements for the convergence of the coefficients of the FIR inverse filter, that is the inverse of the room transfer function, are considered. It is important that the coefficients converge because this is an indication that if a sufficient, finite, number of terms considered the error can be small. It is generally agreed that since the room transfer function is non-minimum phase the inverse filter coefficients will not necessarily converge [46] [44] [39].

In this section it is demonstrated that provided a non-causal filter is acceptable then it is possible to make an exact FIR filter inverse of any FIR filter. (Inverse room acoustic filters have to be inherently non-causal to account to the transit time between the microphone - though this is normally expressed by allowing the desired response to be delayed. Inverse filters in general may require substantially larger delays when dealing with non-minimum phase systems - this does not represent a problem for audio reproduction.) Since any linear system can be expressed as an infinite length FIR filter then it follows from this assertion that any room will have a (non causal) exact inverse FIR filter.

Given that a FIR filter can represent the exact inverse of a room then the calculation of the filter becomes routine. FIR filters, even inverse types, can be produced adaptively using NLMS techniques. As stated before NLMS techniques have many advantages, notably relating to simplicity and stability.

Looking that the inverse of the polynomial $G_{(z^{-1})}$. If the polynomial $G_{(z^{-1})}$ is factorized then the inverse is given by the product of the inverse of the factors:

$$H_{(z^{-1})} = \frac{z^{-\Delta}}{(z^{-1}-r_{g1})(z^{-1}-r_{g2})\cdots} \tag{9.25}$$

$$= z^{-\Delta}\frac{1}{(z^{-1}-r_{g1})}\frac{1}{(z^{-1}-r_{g2})}\cdots \tag{9.26}$$

The factors may be both greater and less than unity and may be complex. If initially only real factors are considered then from the Binomial Theorem the inverse of each of the factors of $G_{(z)}$ is given by:

$$\frac{1}{(z^{-1}-r_{gn})} = \frac{1}{-r_{gn}}\left(1-\frac{z^{-1}}{r_{gn}}\right)^{-1} \tag{9.27}$$

$$= \frac{1}{-r_{gn}}\left[1+\frac{z^{-1}}{r_{gn}}+\frac{z^{-2}}{-r_{gn}^{2}}+\cdots+\frac{z^{-i}}{r_{gn}^{i}}\right] \quad for \quad -1<\frac{z^{-1}}{r_{gn}}<1 \tag{9.28}$$

It is possible to demonstrate the accuracy of this result using the sum of a geometric series:

$$\sum_{i=1}^{\infty} ar^{i-1} = \frac{a}{(1-r)} \tag{9.29}$$

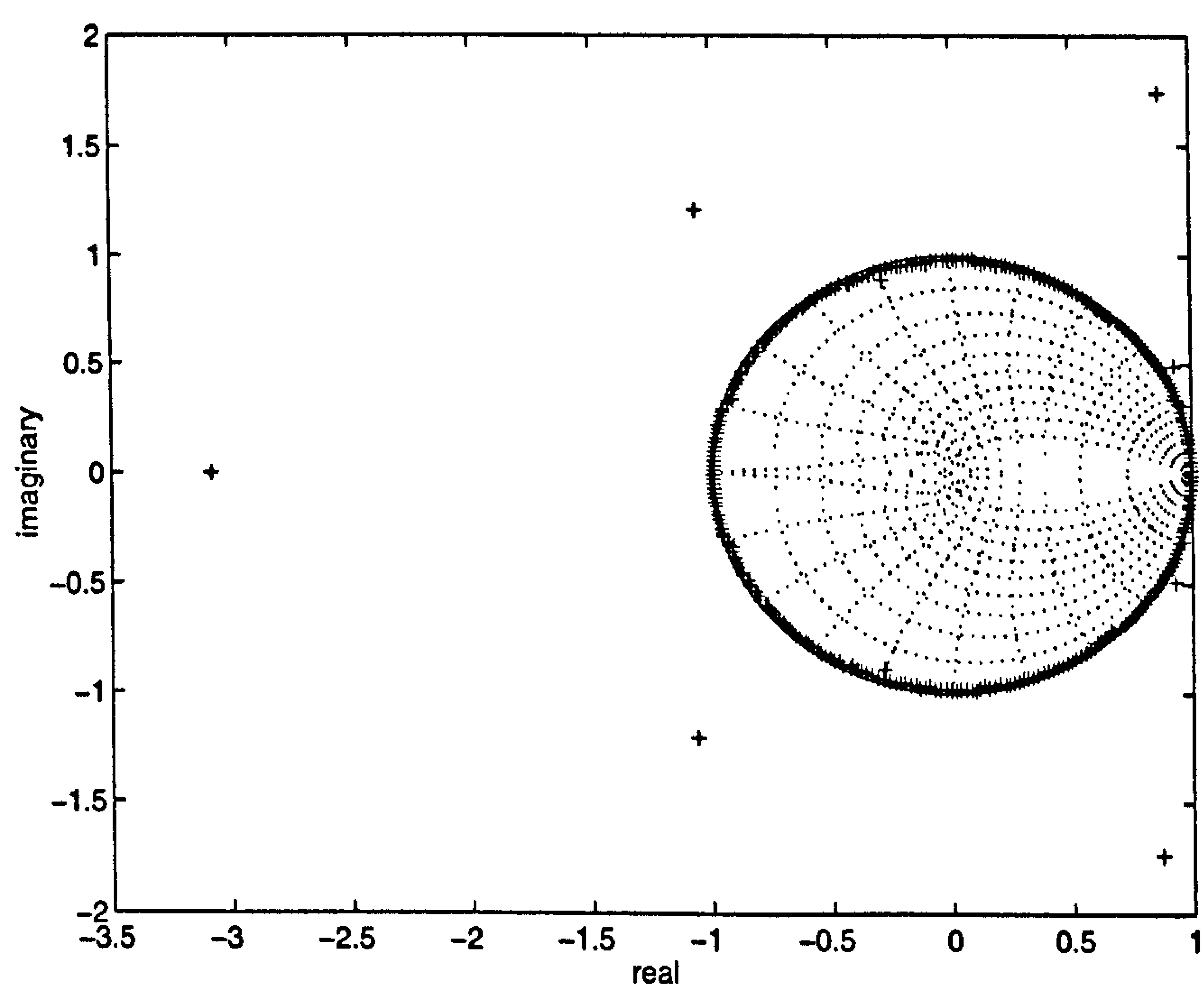
$$\frac{1}{-r_{gn}} \sum_{i=1}^{\infty} \left(\frac{z^{-1}}{r_{gn}}\right)^{i-1} = \frac{1}{-r_{gn}} \frac{1}{(1-\frac{z^{-1}}{r_{gn}})} \tag{9.30}$$

$$= \frac{1}{(z^{-1}-r_{gn})} \tag{9.31}$$

If each of the roots is expanded separately then the final solution is the results of all of the expansions convolved. (Multiplication of polynomials is equivalent to convolution of their coefficients.) This yields the coefficients of the FIR filter that is the exact inverse of the room transfer function.

From inspection of the binomial expansion that provided that $\left(\frac{1}{r_{gn}}\right)^{i} < \left(\frac{1}{r_{gn}}\right)^{i+1}$ the series will converge. This implies that r_{gn} must be greater than unity. This criterion can also be stated as the root must fall outside the unit circle. In practice there are many roots that do not meet this criterion. The graph below is a plot of the roots for a real room transfer function from 2 - 200Hz approximately. (Note that most of the roots are just inside the unit circle.)

Figure 9.5: Roots of $G_{(z^{-1})}$



The convergence criterion can also be demonstrated from the constraint of the binomial expansion, $-1 < \frac{z^{-1}}{r_{g1}} < 1$, since $|z^{-1}| = 1$. This is reasonable since multiplying any function by z^{-1} will not change the magnitude of the function; z^{-1} is a pure time delay and has magnitude unity. Placing this information in the constraint yields: $-1 < \frac{1}{r_{g1}} < 1$ or $\left|\frac{1}{r_{g1}}\right| < 1$.

The next stage is to find a way to deal with roots with a magnitude of less than one. This method of splitting the problem and then dealing with each half separately was inspired by [59]; Tohyama *et al* only dealt with the minimum phase part of the problem after splitting. If the inversion problem is re-stated in factorized form with the factors inside and outside a unit circle grouped.

$$H_{(z)} = z^{-delay} \frac{1}{(z^{-1} - r_{g_{gt}1})(z^{-1} - r_{g_{gt}2}) \cdots} \frac{1}{(z^{-1} - r_{g_{lt}1})(z^{-1} - r_{g_{lt}2}) \cdots} \tag{9.32}$$

Looking again at a single factor of $G_{(z^{-1})}$ which is less than unity, $\frac{1}{z^{-1} - r_{g_{lt}n}}$.

$$\begin{aligned} \frac{1}{(z^{-1} - r_{g_{lt}n})} &= \frac{z}{(1 - r_{g_{lt}n}z)} & (9.33) \\ &= z\,(1 - r_{g_{lt}n}z)^{-1} & (9.34) \end{aligned}$$

expanding this using the binomial theorem:

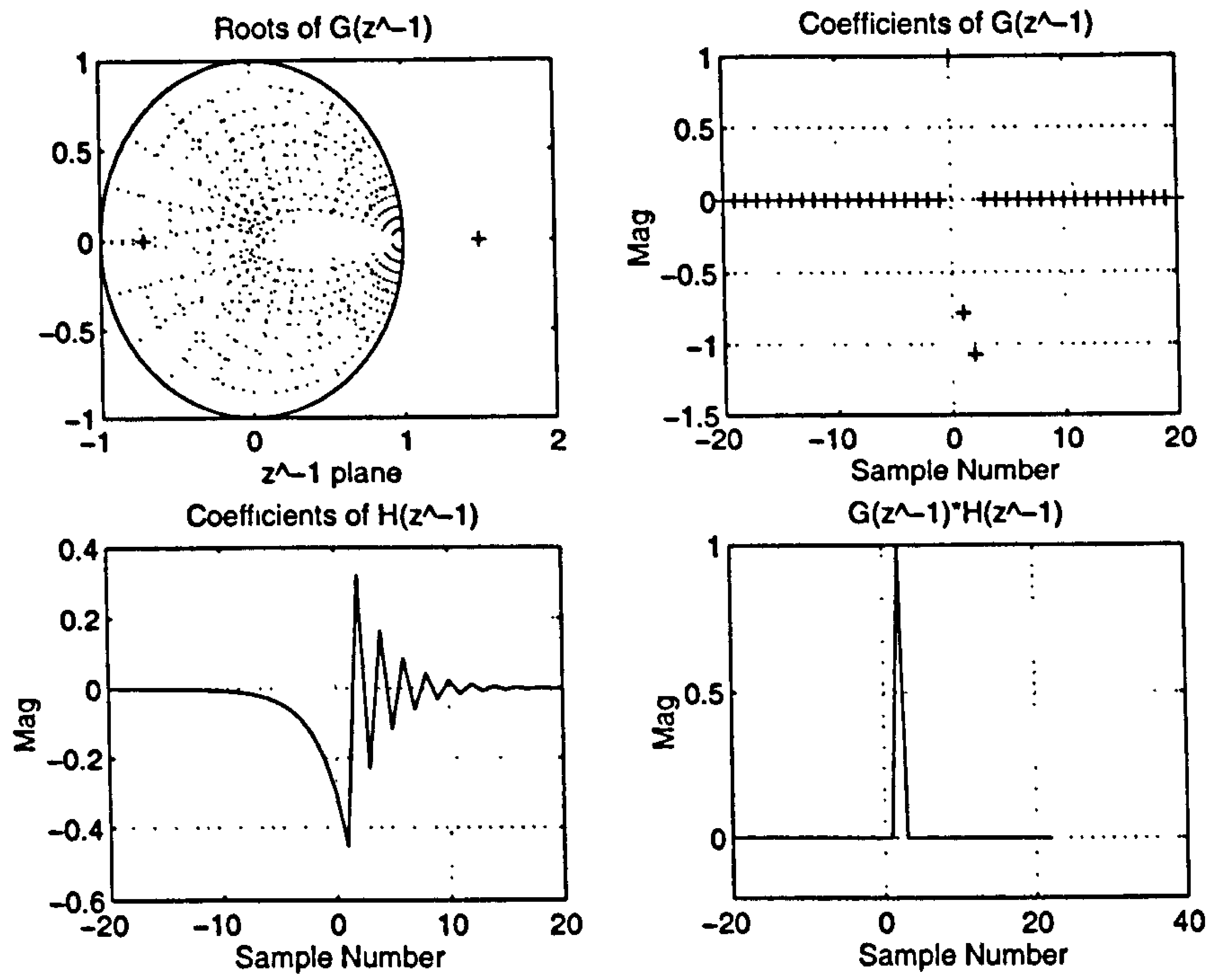
$$z\,(1 - r_{g_{lt}n}z)^{-1} = z\left[1 + r_{g_{lt}n}z + r^2_{g_{lt}n}z^2 + r^3_{g_{lt}n}z^3 + \cdots\right] \qquad for \quad |zr_{g1}| < 1 \tag{9.35}$$

Given that $|z^{-1}| = 1$, as argued previously, then the constraint becomes $|r_{g1}| < 1$. Since this filter is in terms of z rather than z^{-1} it must extend forwards in time. It is not possible to implement this type of filter in a causal system directly. In practice this does not represent a problem since a delay can be included so that a truncated version of this series can be processed. Convolving the series expansions for the inverse of each of the factors of $G_{(z^{-1})}$ will yield $H_{(z^{-1})}$. An example is given below. Note pure delays are a subset of roots less than unity, since they are effectively roots at zero, and thus they can be dealt with without problem. (A pure delay is one of the simplest examples of a non-minimum phase system.)

$$\begin{aligned} G_{(z^{-1})} &= (z^{-1} - a)(z^{-1} - b) \qquad |a| > 1 \quad |b| < 1 & (9.36) \\ H_{(z^{-1})} &= \frac{1}{(z^{-1} - a)} \frac{1}{(z^{-1} - b)} & (9.37) \\ \frac{1}{(z^{-1} - a)} &= \frac{1}{-a}\left[1 + \frac{z^{-1}}{a} + \frac{z^{-2}}{a^2} + \cdots\right] & (9.38) \\ \frac{1}{(z^{-1} - b)} &= z\left[1 + bz + b^2z^2 + \cdots\right] & (9.39) \end{aligned}$$

Given $a = \frac{3}{2}$ and $b = \frac{-5}{7}$ and using Matlab to produce a truncated expansion of the two series the results in figure 9.6 were obtained.

Figure 9.6: Demonstration for Two Roots



It is necessary that the convolution of two of these convergent series is itself convergent (possibly in both directions from the time origin). This problem is problem is considered in detail later.

In addition to roots inside and outside of the unit circle there are complex roots that can occur in either location, but the series of coefficients resulting from their inversion cannot be dealt with by the previous method. The series expansion of the inverse of a complex factor is dealt with in Appendix 9.6. No exact, general proof of convergence was found but it can be demonstrated that for a large number of cases the assertion holds. (Note that a root having a magnitude of exactly unity will not produce a convergent inverse, but this is a singular case in a continuous system and thus can reasonably be neglected. As the roots approach unity more and more coefficients will be required to attain reasonable performance. Providing alternative paths can help reduce the incidence of this type of event and thus significantly reduce the required length of the filter.)

Considering the inverse of the following factor:

$$\frac{1}{z^{-2} - 2az^{-1} + a^2 + b^2} \tag{9.40}$$

Using Matlab to produce the series expansions using the general term discussed in Appendix 9.6 for a large number of values of a and b and a large number of terms of the series expansion.

In the figures 9.7 and 9.8 the values of a and b for which each of the coefficients produced by the series expansion of the inverse of a pair of complex roots is less than those of a pair of real roots with the same magnitude is shown (Demonstrating that every term of a series is less than the equivalent term of a series known to converge is a sufficient demonstration of convergence, see Appendix 9.6 for further details.) The results are predictable, all values of a and b tested which correspond to roots with a magnitude of greater than unity produce apparently convergent series expansions. As explained before it is possible to restate the problem in terms of z instead of z^{-1}, the effect of which is convergence for values of a and b corresponding to roots with a magnitude of less than unity.

Figure 9.7: Convergence for Roots Corresponding to a and b.

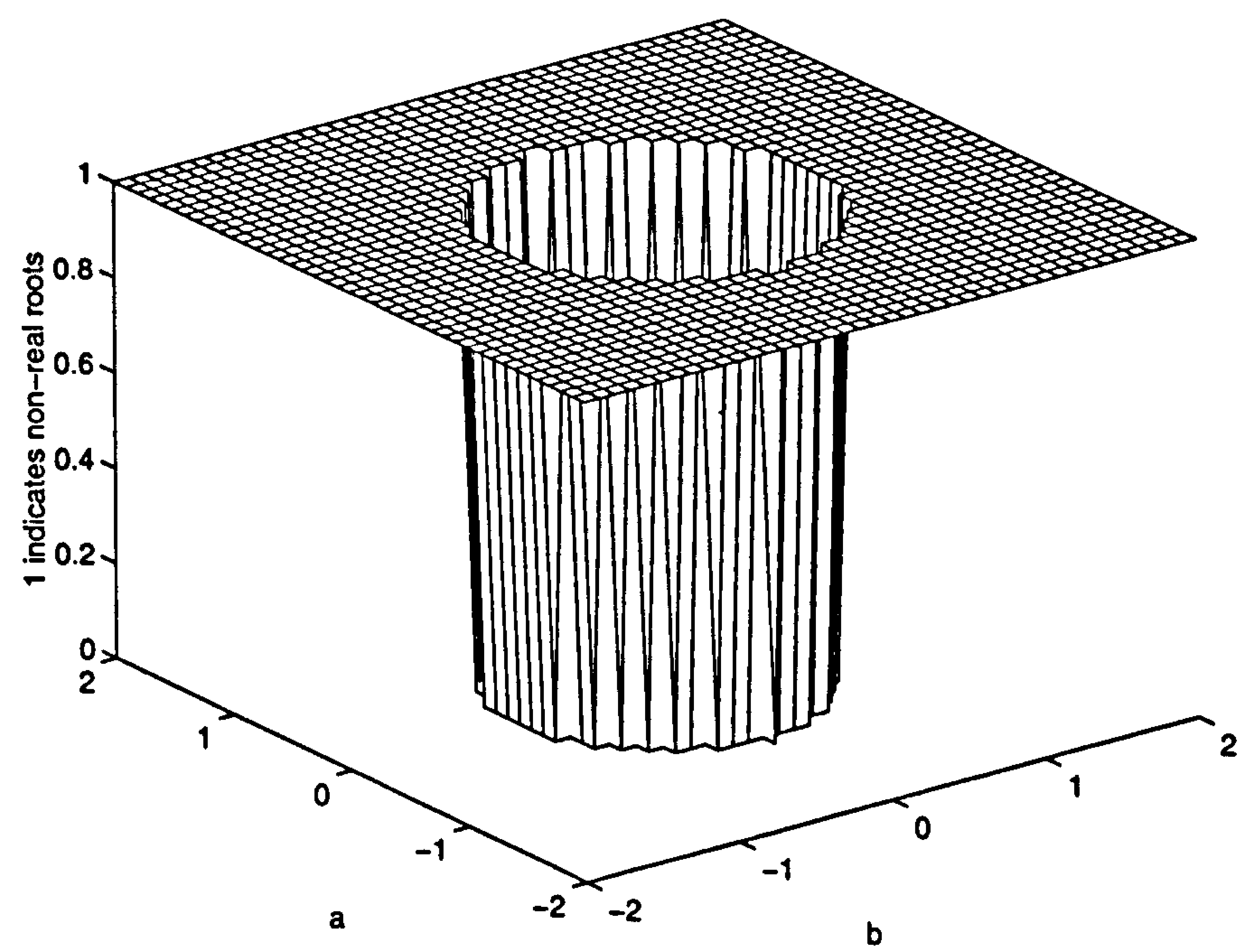
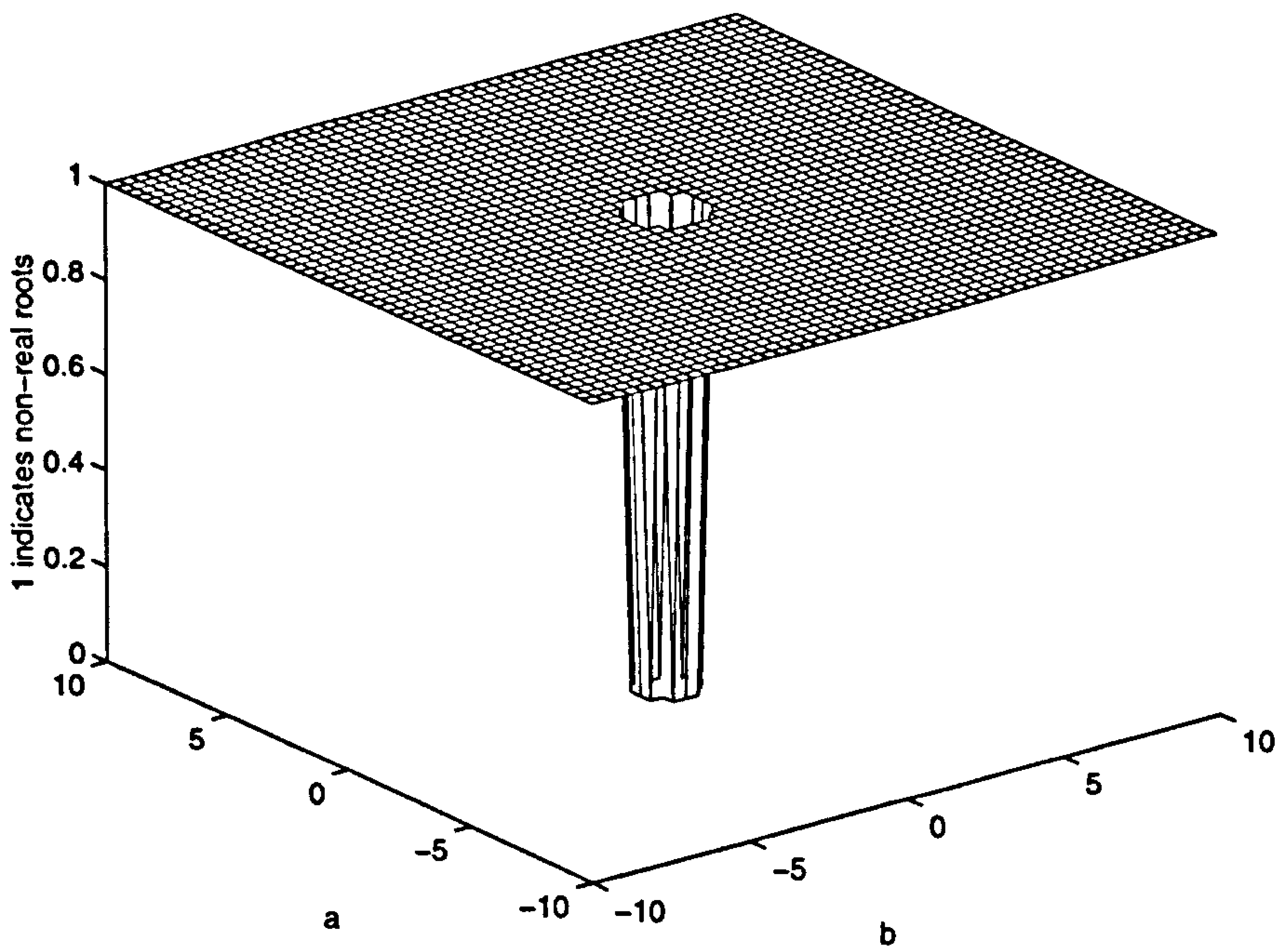


Figure 9.8: Convergence for Roots Corresponding to a and b.



The demonstration that it is possible to produce an exact inversion provided that the filter is allowed to extend both ways to infinity is complete when it is considered that convolving any finite number of convergent series expansions will always yield a filter that converges as it approaches infinity in either direction. A general demonstration in support of this claim is given below, alternative demonstrations may be found in Appendix 9.5:

Taking the general case of a convolution of two convergent series:

$$A = \sum_{n=1}^{\infty} a_n \tag{9.41}$$

$$B = \sum_{n=1}^{\infty} b_n \tag{9.42}$$

$$\tag{9.43}$$

Where A and B are real numbers, ie. the two series converge.

$$A * B = \quad a_1 [b_1 + b_2 + b_3 + b_4 + \cdots + b_n] \tag{9.44}$$

$$+ \quad a_2 [0 + b_1 + b_2 + b_3 + \cdots + b_n] \tag{9.45}$$

$$+ \quad a_3 [0 + 0 + b_1 + b_2 + \cdots + b_n] \tag{9.46}$$

$$+ \quad \vdots \tag{9.47}$$

$$+ \quad a_n [0 + \cdots + b_1 + b_2 + \cdots + b_n] \tag{9.48}$$

Using the definition of B this may be re-written as:

$$
\begin{aligned}
A * B = \quad & a_1 [B] & (9.49)\\
+ \quad & a_2 [B] & (9.50)\\
+ \quad & a_3 [B] & (9.51)\\
+ \quad & \vdots & (9.52)\\
+ \quad & a_n [B] & (9.53)
\end{aligned}
$$

Since a_n is a convergent series use may be made of one of the basic properties of this type of sequence; if x_n is the general term of a convergent series and C a real number then:

$$\sum_{n=1}^{\infty} x_n C = C \sum_{n=1}^{\infty} x_n \tag{9.54}$$

Thus:

$$
\begin{aligned}
A * B &= \sum_{n=1}^{\infty} a_n B & (9.55)\\
&= B \sum_{n=1}^{\infty} a_n & (9.56)\\
&= AB & (9.57)
\end{aligned}
$$

Thus, in the general case, if two series converge then the series created by convolution of their terms will converge also and the sum of this new series will be the product of the sums of the initial series.

While series are being discussed it is interesting to note that with an alternating series it is possible to place a bound the error that is incurred if only the first n terms are considered.

As noted in section 7.2 Nelson *et al* produced an exact formula in [46] which predicted "the number of coefficients required in inverse filters if an exact inverse is to exist."

$$I = \frac{\left[\sum_{l=1}^{L} J_l\right] = L}{M - L} \tag{9.58}$$

where:

I = Length (in samples) of the entire inverse filter vector

J_l = Length (in samples) of the l^{th} transmission path model

L = Number of microphones

M = Number of Loudspeakers

It can be seen that the formula Nelson *et al* proposed makes no reference to the room system in question, being concerned only with the number of reproduction channels and lengths of each filter.

When a single root is inverted to produce a series expansion then is some cases the series will be alternating. If such an alternating series is considered then some indication of the residual error can be obtained. Given that an alternating series $\sum_{n=1}^{\infty}(-1)^{n-1}a_n$ converges then the remainder is given by:

$$R_n \leq c_{n+1} \tag{9.59}$$

The n^{th} term of the expansion of a root is $z^{-n}r_g^{n-1}$ for $|R_g| < 1$ thus as r_g approaches unity the value of the error will become slow to decrease. For this reason it is not possible to set a number of elements that will be required in the filter for a certain accuracy without knowledge of the roots of the room transfer function. This result casts doubt on the result obtained by Nelson *et al* in [46] where Nelson gives an explicit formula for the number of elements required in the filter without recourse to the roots of the room transfer function.

Nelson *et al*'s error probably lies in confusing the room model (which is truncated) with the real system which is infinite. Using a truncated model and multiple channels it is possible to arrange matters such that the matrix that requires inversion is square and thus exact inversion is possible. However, since the room model is a truncated version of the real system the filter will *not* be exact when applied to the real system and thus this potentially beguiling result is practically meaningless. More in-depth consideration of this effect is given in section 7.6.1.

An alternative way of thinking about the problem is to think about the stability criterion for the z^{-1} domain representation. The inverse of a root of the room transfer function expressed in terms of z^{-1} is a pole in the z^{-1} plane, the stability of which is assured provided that its magnitude is greater than unity [15]. Stability and convergence are not interchangeable; if the coefficients of the impulse response of a pole are convergent then the pole is stable but there are some cases, such as the harmonic series, where the values of the coefficients reduce and even tend to zero yet the series is not convergent.

9.3 The Matrix Method

This section is the details the "matrix method" and includes a justification for the results used in the main body of the text. The matrix method is not an adaptive method per-say, it is a way of stating the inversion problem that allows the use of high level mathematics packages such as Matlab or the NAG libraries.

Given that **g** is the a vector representing the room impulse response after appropriate truncation, **h** is the filter and **d** is the desired impulse response where:

$$\mathbf{d} = \begin{bmatrix} d_0 \\ d_1 \\ \vdots \\ d_{L_g+L_h+1} \end{bmatrix} \qquad \mathbf{g} = \begin{bmatrix} g_0 \\ g_1 \\ \vdots \\ g_{L_g} \end{bmatrix} \qquad \mathbf{h} = \begin{bmatrix} h_0 \\ h_1 \\ \vdots \\ h_{L_h} \end{bmatrix} \tag{9.60}$$

and

$$\mathbf{G} = \begin{bmatrix} g_0 & & 0 \\ g_1 & g_0 & \\ \vdots & g_1 & \ddots \\ g_{L_g} & \vdots & g_0 \\ & g_{L_g} & g_1 \\ 0 & \ddots & \vdots \\ & & g_{L_g} \end{bmatrix} \tag{9.61}$$

Then ideally:

$$\mathbf{d} = \mathbf{G}\mathbf{h} \tag{9.62}$$

In practice there will be no exact solution to this problem. If an additional variable is added to the equation it is possible to arrange matters such that the equations are now consistent:

$$\begin{bmatrix} g_0 & & 0 \\ g_1 & g_0 & \\ \vdots & g_1 & \ddots \\ g_{L_g} & \vdots & g_0 \\ & g_{L_g} & g_1 \\ 0 & \ddots & \vdots \\ & & g_{L_g} \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{bmatrix} - \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ \vdots \\ d_m \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ \vdots \\ r_m \end{bmatrix} \tag{9.63}$$

$$\mathbf{Gh} - \mathbf{d} = \mathbf{r} \tag{9.64}$$

The "best" solution to this problem can be obtained by minimizing the sum of the squared errors:

$$\mathbf{r^T r} = \begin{bmatrix} r_1 r_2 \cdots\cdots r_m \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ \vdots \\ r_m \end{bmatrix} \tag{9.65}$$

$$= r_1^2 + r_2^2 + \cdots + r_m^2 \tag{9.66}$$

$$= \left[\mathbf{Gh} - \mathbf{d}\right]^T \left[\mathbf{Gh} - \mathbf{d}\right] \tag{9.67}$$

$$= \mathbf{h^T G^T G h} - \mathbf{h^T G^T d} - \mathbf{d^T G h} + \mathbf{d^T d} \tag{9.68}$$

$$= \mathbf{h^T G^T G h} - 2\mathbf{h^T G^T d} + \mathbf{d^T d} \tag{9.69}$$

$\mathbf{h^T G^T d} = \left[\mathbf{d^T G h}\right]^T$ which equals $\mathbf{d^T G h}$ in this case since all of these are scalars, they have to be since their sum is equal to the sum of the error squared which is also a scalar.

The next step is to examine the partial derivatives of this equation with respect to $h_1, h_2, \cdots$, the minimum error occurs where the gradient is equal to zero for a system with a stable inverse. Considering the general case of h_i:

$$\frac{\partial\left(\mathbf{r^T r}\right)}{\partial h_i} = \frac{\partial}{\partial h_i}\left[\mathbf{h^T G^T G h}\right] - 2\frac{\partial}{\partial h_i}\left[\mathbf{h^T G^T d}\right] \quad \mathit{for}\ \ 0 < i \leq n \tag{9.70}$$

$$= \begin{bmatrix} h_1 \\ \vdots \\ h_{i-1} \\ h_i \\ h_{i+1} \\ \vdots \\ h_n \end{bmatrix}^T \mathbf{G^T G} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1_i \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1_i \\ 0 \\ \vdots \\ 0 \end{bmatrix}^T \mathbf{G^T G} \begin{bmatrix} h_1 \\ \vdots \\ h_{i-1} \\ h_i \\ h_{i+1} \\ \vdots \\ h_n \end{bmatrix} - 2 \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1_i \\ 0 \\ \vdots \\ 0 \end{bmatrix} \mathbf{G^T d} \tag{9.71}$$

Since $\mathbf{G^T G}$ is symmetrical about the leading diagonal:

$$\begin{bmatrix} h_1 \\ \vdots \\ h_{i-1} \\ h_i \\ h_{i+1} \\ \vdots \\ h_n \end{bmatrix}^T \mathbf{G^T G} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1_i \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1_i \\ 0 \\ \vdots \\ 0 \end{bmatrix}^T \mathbf{G^T G} \begin{bmatrix} h_1 \\ \vdots \\ h_{i-1} \\ h_i \\ h_{i+1} \\ \vdots \\ h_n \end{bmatrix} \tag{9.72}$$

Setting the derivative equal to zero in order to find the minimum:

$$\frac{\partial\left(\mathbf{r^T r}\right)}{\partial h_i} = 0 \quad \Rightarrow \quad 2\begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1_i \\ 0 \\ \vdots \\ 0 \end{bmatrix}^T \mathbf{G^T G} \begin{bmatrix} h_1 \\ \vdots \\ h_{i-1} \\ h_i \\ h_{i+1} \\ \vdots \\ h_n \end{bmatrix} = 2\begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1_i \\ 0 \\ \vdots \\ 0 \end{bmatrix}^T \mathbf{G^T d} \tag{9.73}$$

Given that $\mathbf{G^T G h}$ and $\mathbf{H^T d}$ are $n \times 1$ column vectors and that $0 < i \leq n$ limits the values that i can take - i.e the multiplication is conformable - then it is true that all of the elements of the column vector are equal. Thus:

$$\mathbf{G^T G h}_{opt} = \mathbf{G^T d} \tag{9.74}$$

Thus it is reasonable to state that the solution with the least square error to the original problem is given by:

$$\mathbf{h}_{opt} = \left[\mathbf{G^T G}\right]^{-1} \mathbf{G^T d} \tag{9.75}$$

Thus the optimum value of h can be calculated explicitly from knowledge of G and **d**.

An alternative notation is to take the expectation (mean) of the squared error, but this has no effect on the answer. The values of the filter taps can be seen to be independent of each other.

This is an important result since it allows the direct calculation of the optimum setting for **h**. Unfortunately, since no reference is made to the signal x_k this derivation cannot be used to form the basis of an iterative algorithm.

Substituting this value for the solution back into the original equation gives a value for the minimum sum of squared error. (I is the identity matrix.)

$$\mathbf{r^T r} = \mathbf{h^T G^T G h} - 2\mathbf{h^T G^T d} + \mathbf{d^T d} \tag{9.76}$$

$$= \left[\left[\mathbf{G^T G}\right]^{-1}\mathbf{G^T d}\right]^{\mathbf{T}}\mathbf{G^T G}\left[\mathbf{G^T G}\right]^{-1}\mathbf{G^T d} \tag{9.77}$$

$$-2\left[\left[\mathbf{G^T G}\right]^{-1}\mathbf{G^T d}\right]^{\mathbf{T}}\mathbf{G^T d} + \mathbf{d^T d} \tag{9.78}$$

$$= \left[\left[\mathbf{G^T G}\right]^{-1}\mathbf{G^T d}\right]^{\mathbf{T}}\mathbf{I G^T d} - 2\left[\left[\mathbf{G^T G}\right]^{-1}\mathbf{G^T d}\right]^{\mathbf{T}}\mathbf{G^T d} + \mathbf{d^T d} \tag{9.79}$$

$$= \mathbf{d^T d} - \left[\left[\mathbf{G^T G}\right]\mathbf{G^T d}\right]^{\mathbf{T}}\mathbf{G^T d} \tag{9.80}$$

$$= \mathbf{d^T d} - \mathbf{h^T_{opt} G^T d} \tag{9.81}$$

This is a useful result since allows a lower bound to be placed on the Mean Squared Error (MSE) between the optimum filter $\mathbf{h}_{opt}$ and the desired result $\mathbf{d}$.

9.4 Matlab Code To Measure Room Impulse Responses

9.4.1 Top Level Matlab PRBS Measurement Function

```
% Top level m-file for impulse response measurement.
%
% [resp,signal,sigrec] = mat_rpm(filename,stages,
% factor,repeats,cycles,filter,blank)
%
% filename = the root filename used for the recorded and played
% sequences.
% stages = number of stages in the prbs generator.
% factor = the factor by which the sequence is lengthened to.
% repeats = the number of times the prbs sequence appears in the
% sequence.
% cycles = the number of times the sequence is used.
% blank = blank space to precede the sequence, to allow DAC to
% stabilize.
%
% resp = Impulse response of the system under test
% signal = A single cycle of the signal used for measurement
% sig_rec = The mean recorded signal sequence

function [resp,signal,sigrec,sigmod] = mat_rpm(filename,...
          stages,factor,repeats,cycles,filter,blank);

delay = 76;    % An estimate of a delay inherent in the hardware.

% set levels is a function to aid setting the gain of the input stage.

set_levels(1);

% sequence_gen is a function to generate a single cycle of the
```

```
% signal used to measure the impulse response. Sequence is saved
% direct to disk.

[signal,rectime,sigmod]=sequence_gen(filename,stages,factor,repeats,
filter,blank,1);

sigrec = zeros(1,length(signal)*2+1);

% The loop runs for the number of cycles specified. The sequence
% is replayed over the BCE AES interface and the results are recorded
% concurrently to disk. The recorded signal is then loaded from
% disk and the peak value checked. Finally the recorded signal is
% sent to the function recover which takes the recorded signal
% and produces an average recorded sequence.

for cycle = 1 : cycles,

   disp(['Cycle = ' num2str(cycle)]);
   funct='!s56x_rpm -p bceaes -m -b 24 ';
   funct=[funct filename ' ' num2str(rectime)];
   eval(funct);

   loadname=[filename '.rec'];
   sigrec1 = b_load(loadname,1);

   peak = max(abs(sigrec1(length(signal)*factor+blank:...
             length(signal)*(factor*3)+blank)));
   disp(['Peak signal recorded = ' num2str(peak)]);
   if peak>130000
      Warning=['Signal probably clipped.']
   end

   [sigrec1] = recover(signal,sigrec1,factor,repeats,blank,delay,1);
```

```
  sigrec = sigrec+sigrec1;

end

% The recorded signal is normalized and then the impulse response
% is calculated by means of cross correlation. The portion of
% the response corresponding to the system impulse response is
% selected.

sigrec = sigrec./cycles;
signal = signal./65535;
resp = (real(xcorr(signal,sigrec)))./(length(signal)+1);
resp = [resp(length(signal)*2:length(signal)*3-ceil(0.02*length(signal)))
        zeros(ceil(0.02*length(signal)),1)];
       % Last 2 percent blanked to avoid the tail of the next repetition

% The recorded response is plotted.

subplot(211);
plot(resp),title('TIME'),ylabel('mag'),xlabel('Time in samples'),grid;
subplot(212);
mag = 20*log10(abs(fft(resp)));
n = max(size(resp))/2;
f = 0.5*48000*(0:n-1)/(n*factor);
semilogx(f,mag(1:n)),title('FREQUENCY'),ylabel('dB'),
grid,
xlabel('Frequency in Hz');
```

9.4.2 Called Function - Set Levels

```
% File to see if the gain is about right for the data collection.
```

```
%
% []=set_levels(chans)

function []=set_levels(chans)

string = input('Do you wish to set the levels?','s');
if string == '',
  string = 'n';
end
while string(1,1)=='y',
   if chans==1,
      !s56x_rpm -p bceaes -b 24 -m  /sw10/round/AES_m_files/level 2;
      level = b_load('/sw10/round/AES_m_files/level.rec',1);
      peak = max(abs(level(1000:10000)))
      if peak>130000
         disp('Warning - Signal probably clipped.');
      end
   elseif chans==2,
      !s56x_rpm -p bceaes -b 24  /sw10/round/AES_m_files/level2 2;
      level = b_load('/sw10/round/AES_m_files/level2.rec',2);
      peak_a = max(abs(level(1,1000:10000)))
      peak_b = max(abs(level(2,1000:10000)))
      if max([peak_a peak_b])>130000
         disp('Warning - Signal probably clipped.');
      end
   end
   string = input('Do you wish to set the levels?','s');
   if string == '',
     string = 'n';
   end
end
```

9.4.3 Called Function - Sequence Generation

```
% Function to test if a new sequence is required and if it is
% then to generate it from the prbs sequences.
%
% [signal]=sequence_gen(filename,stages,factor,repeats,filtered
% ,blank,chans);
%
%  filename = the root filename used for the recorded and played
% sequences.
%  stages = number of stages in the prbs generator
%  factor = the factor by which the sequence is to be lengthened.
%  repeats = the number of times the measurement is to be taken.
%  filtered = true/false 1->true
%  blank = blank space to precede the sequence, to allow DAC to
% stabilise.
%  chans = number of channels to be recorded.

function [signal,rectime,sigmod]=sequence_gen(filename,stages,...
factor,repeats,filter,blank,chans);

load /sw10/round/AES_m_files/prevSet

% Load a PRBS signal of the appropriate length (pre-generated to
% save time.)

load(['prbs',num2str(stages)]);

% The initial test is to see if the user requirements have changed.
% A new sequence is only generated if they have in order to save time.

if (prevStages~=stages|...
```

```
    prevFactor~=factor|...
    prevRepeats~=repeats|...
    prevLenFilter~=length(filter)|...
    prevBlank~=blank|...
    prevChans~=chans|...
    strcmp(prevFilename,filename)~=1),

disp('New sequence being constructed');

% The signal is interpolated to reduce effective bandwidth if the
% entire bandwidth is not required. (It is not possible to set the
% sample rate of the BCE AES equipment directly.)

if (factor~=1)
   sig1 = interp(signal,factor);
   sig1 = sig1.*2*(max(signal)/max(sig1));
else
   sig1 = signal;
end

sigmod = zeros(length(sig1)*factor*(repeats*2+2)+blank,1);

% Generate a signal that is made up of the basic sequence repeated
% repeats times.

for i=1:repeats+1

    sigmod(length(sig1)*(i-1)*2+1+blank:length(sig1)*i*2+blank)...
          = [sig1
             sig1];
end

% Prefilter the test signal if a filter has been specified.
```

```
    if (filter ~=0)
       sigmod = conv(sigmod,filter);
       sigmod = sigmod.*(2*max(signal)/max(sigmod));
       % Following lines remove the system delay inherent in the filter.
       % [void,loc]=max(filter);
       % sigmod = sigmod(loc:length(sigmod)-length(filter)+loc);
    end

    % two channel handling if required

    if chans==2,
        sigmod=[sigmod sigmod]';
    end

    % Set rectime, a variable ultimately used by the BCE AES system.

    rectime = ceil(max(size(sigmod))/48000);

    % save the test signal to disk.

    savename = [filename '.play'];
    b_save(savename,sigmod);
    %save diagnostic sigmod
end

% Save the setup used to create the test signal so that the software
% can detect if the requirements have changed.

prevStages = stages;
prevFactor = factor;
prevRepeats = repeats;
prevLenFilter = length(filter);
prevBlank = blank;
prevFilename = filename;
```

```
prevChans=chans;
save /sw10/round/AES_m_files/prevSet prevStages prevFactor ...
   prevRepeats prevLenFilter prevBlank prevFilename prevChans rectime;
```

9.4.4 Called Function - Recover Signal

```
% Function to recover the average recorded sequence from the
% recorded signal.
%
% First the average recorded sequence is calculated, then it is
% decimated if required and finally it is returned to be
% cross-correlated with the original signal to give the impulse
% response.
%
% The delay is a bug (feature) of the hardware that should be
% taken into account. Done by looking delay samples early for
% arrival of data.
%
% function [sigrec] = recover(signal,sigin,factor,repeats,blank,
% delay,chans);

function [sigrec] = recover(signal,sigin,factor,repeats,blank,...
delay,chans)

% Filter removed since its effects were not as predicted.
% [B,A]=butter(4,0.0005,'high');  % 16Hz approx
% sigin = filtfilt(B,A,sigin);  % zero phase filter
```

```
sigrec = zeros(chans,length(signal)*2*factor+1);

% A loop to take the correct segments of the recorded signal
% and add them to produce the sequence used for cross-correlation.

for i=1:repeats
    sigrec =  sigrec + ...
      sigin(1:chans,length(signal)*factor*(2*i-1)+blank+delay:...
                      length(signal)*factor*(2*i+1)+blank+delay);

end

% decimate the recorded sequence if the transmitted sequence was
% interpolated.

if factor~=1
   for i=1:chans
      sigrec(i,1:length(signal)*2+1)=decimate(sigrec(i,:),factor);
   end
   sigrec = sigrec(:,1:length(signal)*2+1);
end

% Remove any DC offset that exists in the recorded sequence and
% do some preliminary normalization.

for i=1:chans
   sigrec(i,:) = sigrec(i,:)-mean(sigrec(i,:));
   sigrec(i,:) = sigrec(i,:)./(65535*repeats);
end
```

9.4.5 Called Functions - Save and Load Data Functions

```
% Save data from Matlab into a Berkeley readable file.
%
% format b_save('filename',data);

function []=b_save(filename,data);

[fileid,message] = fopen(filename,'w+');
if message != []
    message
    break;
end
count = fwrite(fileid,data,'int24');
fclose(fileid);
```

```
% Load the data stored in a Berkeley save file to Matlab

function [data]=b_load(filename,chans);

[fileid,message] = fopen(filename,'r');
if message != []
    message
    break;
end
[data,count] = fread(fileid,[chans,inf],'int24');
fclose(fileid);
[a,b]=size(data);
if a>b,
```

```
    data=data';
end
```

9.5 Non-General Demonstrations that the Result of Convolving Convergent Series is Convergent

In this section non-general demonstrations that the coefficients resulting from a convolution of convergent series are themselves convergent are examined with the intention of producing a general demonstration. The analysis proved difficult to generalize sufficiently and an alternative method was used in the main text. These section is included in spite of failing to meet its aim since the methods used here are possibly more rigorous.

Considering the case where the two series result from the inversion of equal positive roots with magnitude of less than unity. Each root may be inverted and this produces the following sequence:

$$z(1-rz)^{-1} = z + rz^2 + r^2 z^3 + \cdots \tag{9.82}$$

If the coefficients of the various powers of z are considered then it is possible to make a test for the convergence of this series. Since the series resulted from the application of the binomial theorem it is to be expected that it will converge; the demonstration of this property is included as an example of some of the methods that may be used to demonstrate convergence. The series will be of the form:

$$0 + 1 + r + r^2 + \cdots \tag{9.83}$$

Since any finite number of terms may be removed from a series without affecting its convergent properties the preceding zero will be dropped. The resulting series may be represented as

$$\sum_{n=1}^{\infty} r^{n-1} \tag{9.84}$$

It can clearly be seen that provided that the magnitude of r is less than unity then:

$$\lim_{n\to\infty} r^{n-1} = 0 \tag{9.85}$$

This is a necessary but insufficient condition when demonstrating the convergence of a series. (A good demonstration example of a series that satisfies this condition, but is not convergent, is the harmonic series, $\sum_{n=1}^{\infty} 1/n$ which does not converge.) If d'Alembert's ratio test is employed [9], however, it is possible to prove convergence. "If $\sum_{n=1}^{\infty} a_n$ is a series of positive terms and if $\lim_{n\to\infty} a_{n+1}/a_n = l$ then the series is convergent for $l < 1$, divergent for $l > 1$ and for the case of $l = 1$ no conclusion may be drawn." Since:

$$\lim_{n\to\infty} r^n / r^{n-1} < 1 \tag{9.86}$$

The series $1 + r + r^2 \cdots$ is convergent for the magnitude of r less than unity. If r is allowed to be negative and the series is expressed as:

$$\sum_{n=1}^{\infty} (-1)^{n-1} |r|^{n-1} \tag{9.87}$$

Then convergence can still be guaranteed by applying Leibnitz' test [9] which requires that $|r|^{n-1} \geq |r|^n$ for $n = 1, 2, \cdots$ and $\lim_{n\to\infty} |r|^{n-1} = 0$. Both of these constraints are obviously satisfied for $|r| < 1$. Thus the constraints of the binomial expansion as applied to this problem have been demonstrated as accurate.

Returning to the result of convolving two of these identical series. In the final system each of the terms resulting from the convolution of the expansions of the inverse of each of the roots will be the coefficients of the inverse filter. It is important to determine if, provided that each of the series that are convolved to manufacture it converge, the terms of this final filter will definitely converge towards infinity in either direction. Provided that the final series converges then it is possible to say that the error term in least squares approximation will converge for a sufficiently large number of terms in the inverse filter. If, initially, the simple case of the convolution of two series resulting from the expansion of two identical roots whose magnitude is less than unity the result of the convolution will be:

$$[1] + [2r] + [2r^2 + r^2] + [2r^3 + 2r^3] + [2r^4 + 2r^4 + r^4] + \cdots = 1 + 2r + 3r^2 + 4r^3 + \cdots \tag{9.88}$$

$$= \sum_{n=1}^{\infty} n r^{n-1} \tag{9.89}$$

Since $\sum_{n=1}^{\infty} n$ is unbounded then the series product rule cannot be applied.

Considering the n^{th} term.

$$\lim_{n\to\infty} n r^{n-1} \tag{9.90}$$

The limit of this term is not obvious. If n were redefined temporarily as a real number then the application of L'Hospital's rule [9] (using $\frac{d}{dx}(a^x) = a^x \ln x$) yields:

$$\lim_{n\to\infty} n r^{n-1} = \lim_{n\to\infty} \frac{\frac{d}{dn}(n)}{\frac{d}{dn}(r^{1-n})} \tag{9.91}$$

$$= \lim_{n\to\infty} \frac{1}{r^{-n} + r r^{-n \ln r}} \tag{9.92}$$

$$= 0 \tag{9.93}$$

Thus convergence is possible. Depending on the value of r it can be seen that the series will either be alternating or will have all positive terms. The proof of convergence for either is very similar, the alternating sign is handled independently from the coefficient as before.

Leibnitz's test [9](for alternating series) has slightly more stringent conditions and thus if convergence can be demonstrated for this case then the other will follow. Leibnitz requires that the coefficients are monotonic decreasing and that as $n \rightarrow \infty$ the coefficient tends to zero. The latter has already been shown to be satisfied in this case.

The coefficients resulting from convolution cannot be guaranteed to be monotonic decreasing, there is a maximum at a small value of n, though after that the coefficients do decrease as required. The position of the maximum can be determined if n is temporarily considered real.

$$\begin{aligned} \frac{d}{dn} n r^{n-1} &= r^{n-1} + n(r^{n-1} \ln r).1 && (9.94) \\ &= r^{n-1}(1 + n \ln r) && (9.95) \\ & && (9.96) \end{aligned}$$

Setting the gradient equal to zero:

$$\begin{aligned} -n \ln r &= 1 && (9.97) \\ n &= \frac{-1}{\ln r} && (9.98) \end{aligned}$$

Returning n to natural numbers will be carried out using *ceil*, an operator that returns the next natural number larger than the real number it operates on. (This was chosen to *ensure* that n was larger than the value at which the maximum occurred, it is not important if n is slightly larger than required.) It can be seen that n will be positive and finite for r less than unity. Leibnitz' test is now applied to the series after the maximum and the terms before are discarded; this is acceptable since only a finite number of terms are lost.

$$\begin{aligned} ceil\left(m - \frac{1}{\ln r}\right) r^{ceil\left(m-1-\frac{1}{\ln r}\right)} &\geq ceil\left(m + 1 - \frac{1}{\ln r}\right) r^{ceil\left(m-\frac{1}{\ln r}\right)} && (9.99) \\ &\geq r\left[ceil\left(m - \frac{1}{\ln r}\right) r^{ceil\left(m-1-\frac{1}{\ln r}\right)} + r^{ceil\left(m-1-\frac{1}{\ln r}\right)}\right] && (9.100) \\ ceil\left(m + \frac{-1}{\ln r}\right) &\geq r\left[ceil\left(m + 1 + \frac{-1}{\ln r}\right)\right] && (9.101) \\ 1 &\geq \frac{r\left[ceil\left(m + 1 + \frac{-1}{\ln r}\right)\right]}{ceil\left(m + \frac{-1}{\ln r}\right)} && (9.102) \end{aligned}$$

Returning to real numbers to allow the application of L'Hospital's rule since the limit of this function is not obvious.

$$\lim_{r \rightarrow 1} \left[\frac{r\left[m + 1 + \frac{-1}{\ln r}\right]}{m + \frac{-1}{\ln r}}\right] = \lim_{r \rightarrow 1} \left[\frac{m + 1 + \frac{d}{dr}\left(\frac{-1}{\ln r}\right)}{\frac{d}{dr}\left(\frac{-1}{\ln r}\right)}\right] \quad (9.103)$$

$$= \lim_{r \to 1} \left[\frac{\frac{d^2}{dr^2}\left(\frac{-1}{\ln r}\right)}{\frac{d^2}{dr^2}\left(\frac{-1}{\ln r}\right)} \right] \quad (9.104)$$

$$= 1 \quad (9.105)$$

Thus, for this very limited case, the series resulting from a convolution has been demonstrated to converge.

9.6 Partial Demonstration of the Convergence of the Coefficients Produced by Inverting a Complex Root

In this appendix a partial demonstration of the convergence of the series of coefficients produced by the inversion of a complex root is examined. The demonstration is incomplete but the method appears promising, even though the results are not particularly tractable.

A method that may be considered for dealing with complex poles is polynomial division. It is may be demonstrated with specific examples that provided the magnitude of the complex roots does not exceed unity then polynomial division (deconvolution) will yield a convergent series. If similar techniques are used to those suggested in the previous section and the roots are expressed in terms of z instead of z^{-1} then complex roots with a magnitude of greater than unity can be dealt with. A general demonstration is more difficult.

Attempts to find a general series expansion that will deal with the inversion of a pair of complex roots were unsuccessful. Thus algebraic polynomial division probably represents one of the few methods that might achieve a general result.

If we take a quadratic as the product of a pair of complex roots in the form:

$$\left(z^{-1} - (a+bi)\right)\left(z^{-1} - (a-bi)\right) = z^{-2} - 2az^{-1} + (a^2+b^2) \qquad (9.106)$$

$$= z^{-2} + cz^{-1} + d \qquad (9.107)$$

This implies that:

$$c = -2a \qquad (9.108)$$

$$d = a^2 + b^2 \qquad (9.109)$$

Since this analysis will only be dealing with quadratics whose roots are imaginary it is possible to say that:

$$c^2 \geq 4d \qquad or \qquad (9.110)$$

$$(-2a)^2 \geq 4(a^2+b^2) \qquad (9.111)$$

The area of interest is shown graphically in figures 9.9 and 9.10.

Figure 9.9: Range of Values of a and b for Complex Roots

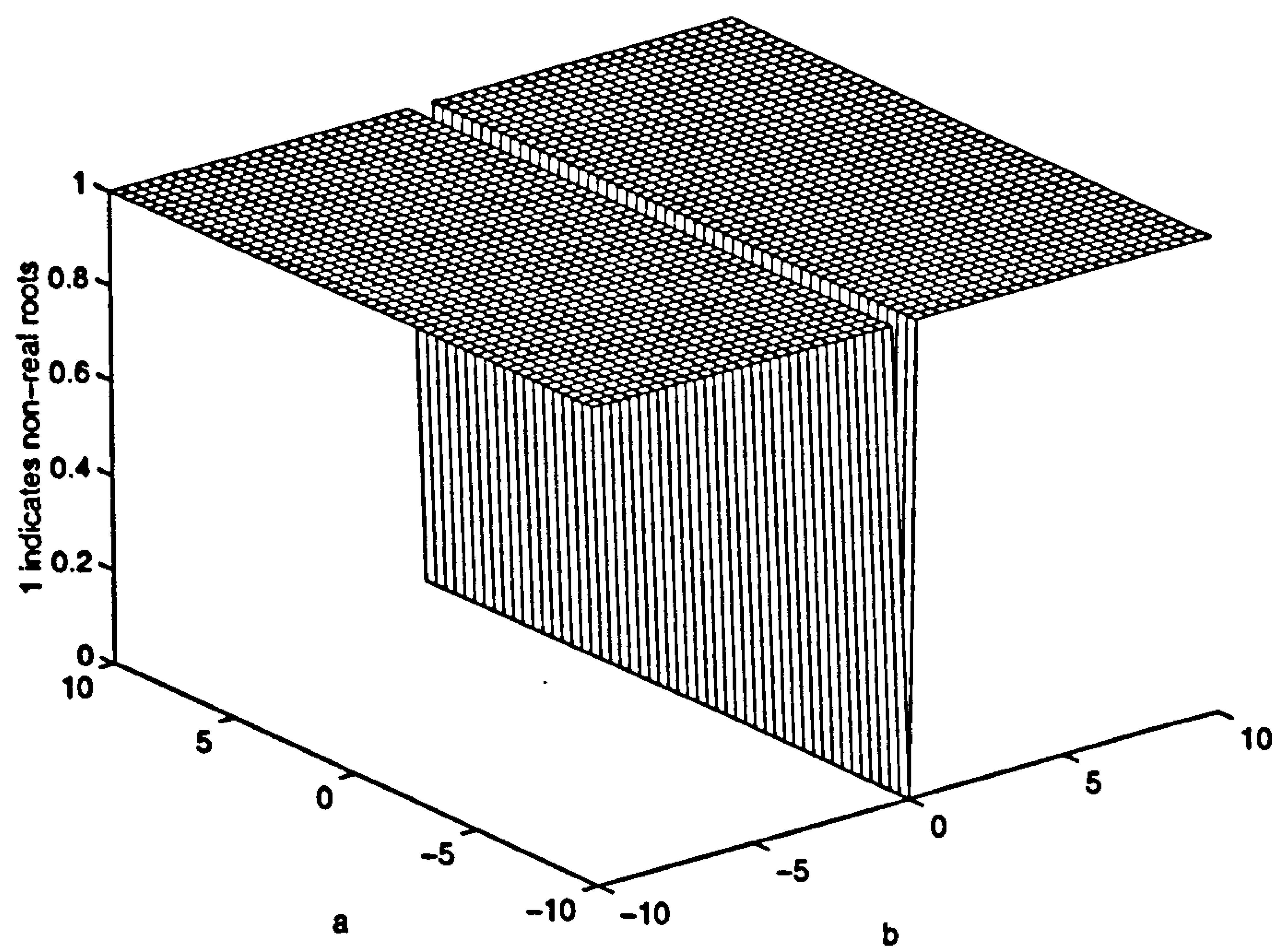
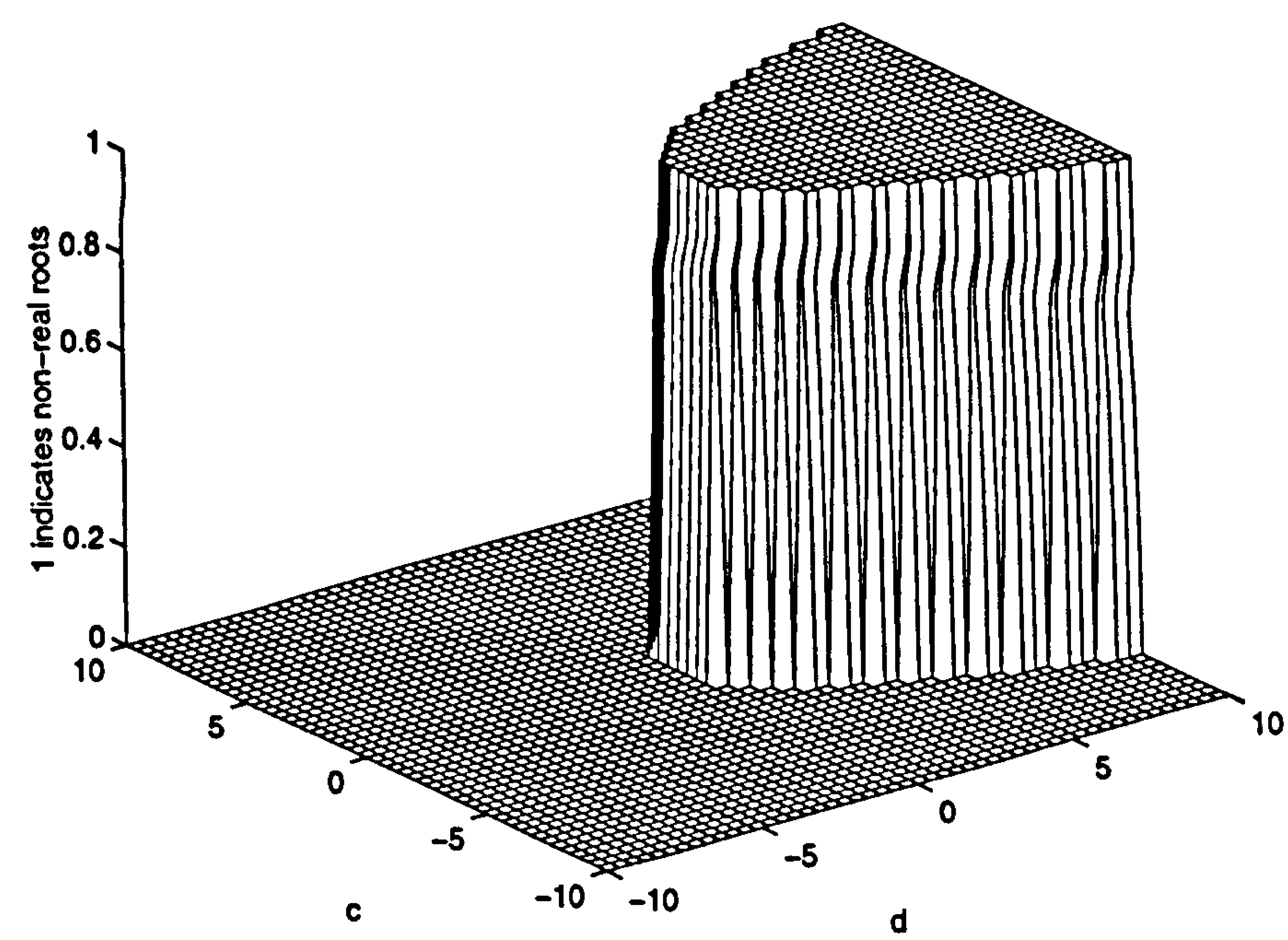


Figure 9.10: Range of Values of c and d for Complex Roots



Calculating the exact inverse of this quadratic is possible by making use of polynomial division. The method is very similar to long division only the result obtained is general. Note that it has been assumed that d is non-zero in order to carry out this process.

$$\frac{1}{z^{-2}+cz^{-1}+d} = \left[\frac{1}{d}\right] \tag{9.112}$$

$$+ \left[-\frac{c}{d^2}\right] z^{-1} \quad (9.113)$$

$$+ \left[\frac{c^2}{d^3} - \frac{1}{d^2}\right] z^{-2} \quad (9.114)$$

$$+ \left[-\frac{c^3}{d^4} + \frac{2c}{d^3}\right] z^{-3} \quad (9.115)$$

$$+ \left[\frac{c^4}{d^5} - \frac{3c^2}{d^4} + \frac{1}{d^3}\right] z^{-4} \quad (9.116)$$

$$+ \left[-\frac{c^5}{d^6} + \frac{4c^3}{d^5} - \frac{3c}{d^4}\right] z^{-5} \quad (9.117)$$

$$+ \left[\frac{c^6}{d^7} - \frac{5c^4}{d^6} + \frac{6c^2}{d^5} - \frac{1}{d^4}\right] z^{-6} \quad (9.118)$$

$$+ \left[-\frac{c^7}{8} + \frac{6c^5}{d^7} - \frac{10c^3}{d^6} + \frac{4c}{d^5}\right] z^{-7} \quad (9.119)$$

$$+ \left[\frac{c^8}{d^9} - \frac{7c^6}{d^8} + \frac{15c^4}{d^7} - \frac{10c^2}{d^6} + \frac{1}{d^5}\right] z^{-8} \quad (9.120)$$

$$+ \left[-\frac{c^9}{d^{10}} + \frac{8c^7}{d^9} - \frac{21c^5}{d^8} + \frac{20c^3}{d^7} - \frac{5c}{d^6}\right] z^{-9} \quad (9.121)$$

$$+ \left[\frac{c^{10}}{d^{11}} - \frac{9c^8}{d^{10}} + \frac{28c^6}{d^9} - \frac{35c^4}{d^8} + \frac{15c^2}{d^7} - \frac{1}{d^6}\right] z^{-10} \quad (9.122)$$

Noting the correspondence between the numeric multiplier in each of the terms which make up the coefficients and Pascal's triangle a general term may be proposed. If the coefficient number (the rows above) is n, where n is a natural number starting from one for the first term and if the term number (the columns above) is k, where k is another natural number starting from one then the expansion above may be expressed as:

$$\frac{1}{z^{-2} + z^{-1} + d} = \sum_{n=1}^{\infty} \sum_{k=1}^{k=n} \left(\begin{array}{c} n-k \\ k-1 \end{array} \right) \frac{c^{n+1-2k}}{d^{n-k+1}} (-1)^{n+k} \quad (9.123)$$

The extra terms included by only limiting the inner sum to $k = n$ are all zero by the definition of Binomial coefficients and thus have no effect on the final sum.

It can be appreciated that for $c = 0$ and d negative and of magnitude greater than unity the series made by the coefficients will converge. If d is allowed positive then the resulting series will alternate with terms of the same magnitude to the equivalent value of d negative. Using Leibnitz' test this series may also be demonstrated convergent. As before it is possible to restate the problem in terms of z instead of z^{-1} and it can be seen that for this case the series made will converge for values of d less than unity. As before for $d = 1$ there is a problem but in a real system the value of d will never be exactly 1.

Demonstrating that the series is convergent in the general case, for c, d limited as indicated before, is more difficult. (The series obviously will not be convergent for $c = 0$ but this is not inside the limits of interest.) Due to the intractability of the general term it seems unlikely that any complete demonstration of the convergent properties of series made by the coefficients of the expansion of a complex pair of roots.

One method of demonstrating series convergence is to prove that each term of the series is less than or equal to the corresponding term of a series that is known to be convergent. (Let $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ be two series of non-negative terms. If there is a natural number N such that $a_n \leq b_n$ for all $n \geq N$, then the convergence of $\sum_{n=1}^{\infty} a_n$ follows from that of $\sum_{n=1}^{\infty} b_n$, and the divergence of $\sum_{n=1}^{\infty} b_n$ from that of $\sum_{n=1}^{\infty} a_n$ [9].) If the magnitude of the terms of the series under test is taken then passing this test demonstrates absolute convergence, of which convergence is a sub-set.

If it is assumed that the series will only converge for the magnitude of the imaginary roots greater than unity then the area of interest is given in the figures 9.11 and 9.12. For the case of roots of less than unity then the problem may be re-expressed in terms of z instead of z^{-1} where the roots will know be greater than unity.

Figure 9.11: a and b Expected to Produce a Convergent Series

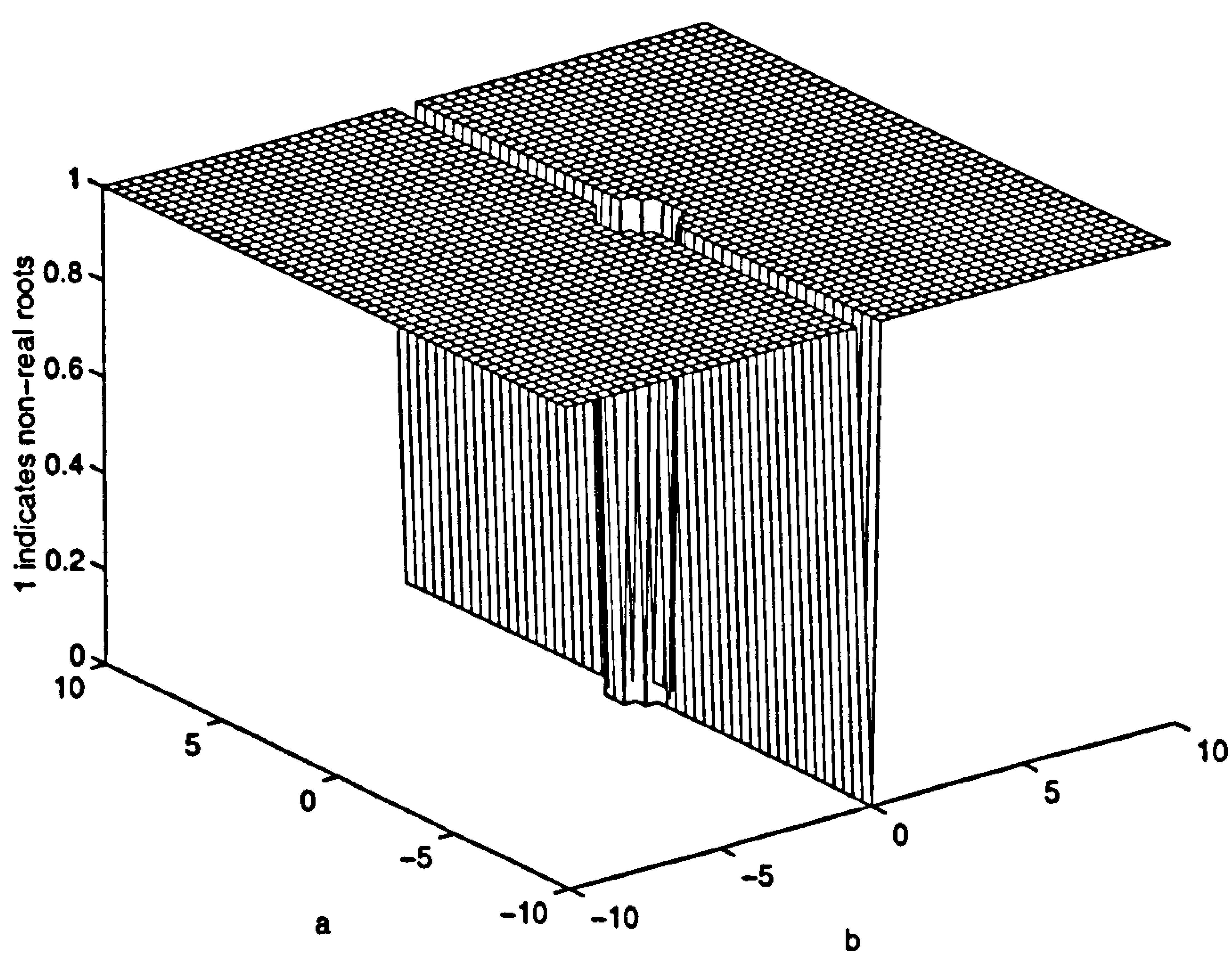
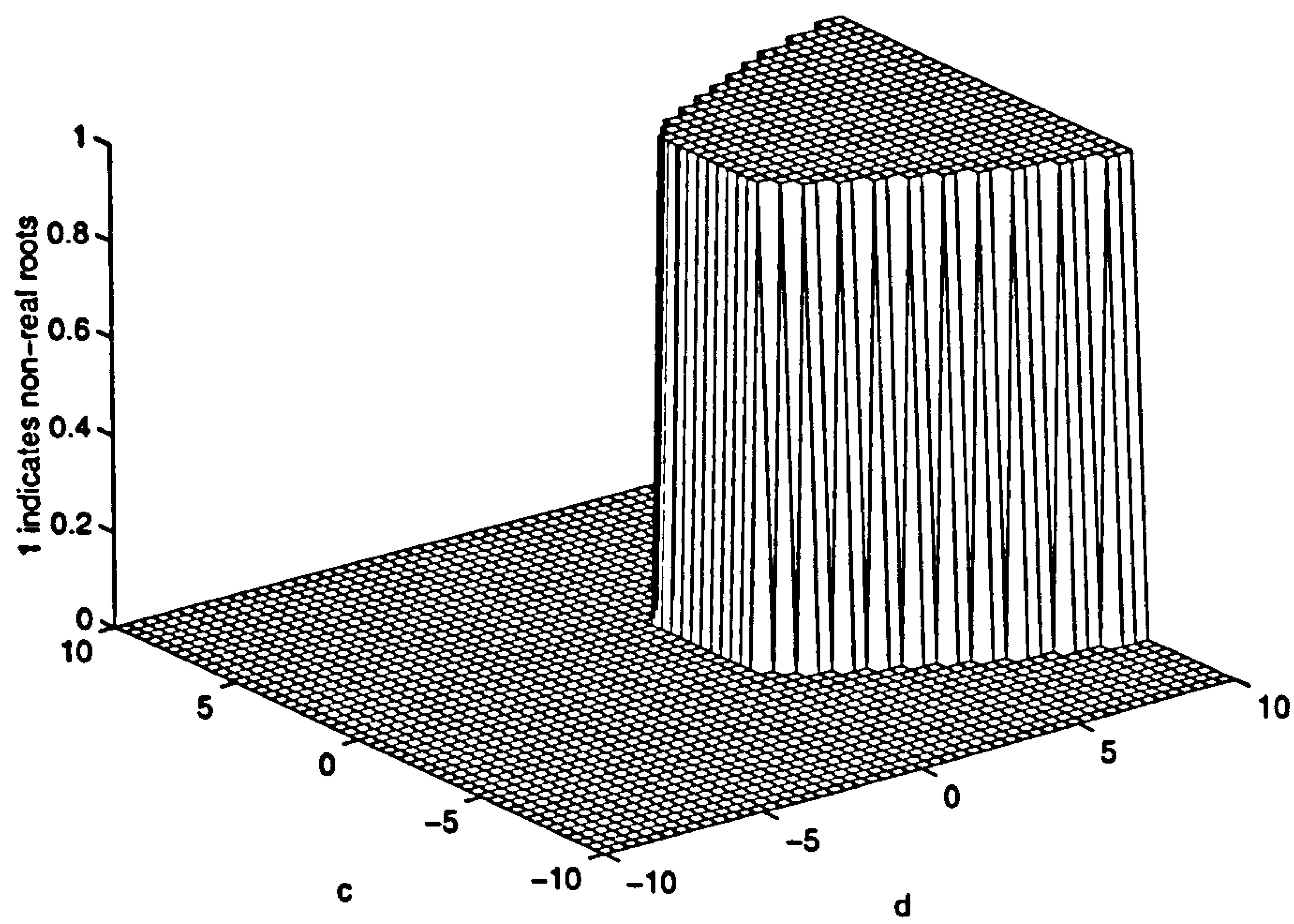


Figure 9.12: c and d Expected to Produce a Convergent Series



If the division is re-expressed in terms of a and b since the valid area is simpler and a series is constructed using the binomial expansions of two roots equal to the magnitude of $a+bi$ convolved.

$$\frac{1}{z^{-2}-2az^{-1}+a^2+b^2} = \left[\frac{1}{a^2+b^2}\right] \tag{9.124}$$

$$+ \left[\frac{2a}{(a^2+b^2)^2}\right] z^{-1} \tag{9.125}$$

$$+ \left[\frac{4a^2}{(a^2+b^2)^3} - \frac{1}{(a^2+b^2)^2}\right] z^{-2} \tag{9.126}$$

$$+ \left[\frac{8a^3}{(a^2+b^2)^4} - \frac{4a}{(a^2+b^2)^3}\right] z^{-3} \tag{9.127}$$

$$+ \left[\frac{16a^4}{(a^2+b^2)^5} - \frac{12a^2}{(a^2+b^2)^4} + \frac{1}{(a^2+b^2)^3}\right] z^{-4} \tag{9.128}$$

$$+ \left[\frac{32a^5}{(a^2+b^2)^6} - \frac{32a^3}{(a^2+b^2)^5} + \frac{6a}{(a^2+b^2)^4}\right] z^{-5} \tag{9.129}$$

$$+ \cdots \tag{9.130}$$

And

$$\frac{1}{z^{-1}-\sqrt{(a^2+b^2)}}\frac{1}{z^{-1}-\sqrt{(a^2+b^2)}} = \frac{1}{(a^2+b^2)} \tag{9.131}$$

$$+ \frac{2}{(a^2+b^2)^{\frac{3}{2}}} z^{-1} \tag{9.132}$$

$$+ \quad \frac{3}{(a^2+b^2)^2}z^{-2} \tag{9.133}$$

$$+ \quad \frac{4}{(a^2+b^2)^{\frac{5}{2}}}z^{-3} \tag{9.134}$$

$$+ \quad \cdots \tag{9.135}$$

Considering coefficients of z^0 it is easy to see that:

$$\frac{1}{(a^2+b^2)} \geq \frac{1}{(a^2+b^2)} \tag{9.136}$$

Considering coefficients of z^{-1}, if:

$$\frac{2}{(a^2+b^2)^{\frac{3}{2}}} \geq \frac{2a}{(a^2+b^2)^2} \tag{9.137}$$

then,

$$\frac{4}{(a^2+b^2)^3} \quad \geq \quad \frac{4a^2}{(a^2+b^2)^4} \tag{9.138}$$

$$\frac{4(a^2+b^2)}{(a^2+b^2)^4} \quad \geq \quad \frac{4a^2}{(a^2+b^2)^4} \tag{9.139}$$

$$\tag{9.140}$$

This is true for a, b real, as they were defined.

Considering coefficients of z^{-2}, if:

$$\frac{3}{(a^2+b^2)^2} \geq \frac{4a^2}{(a^2+b^2)^3} - \frac{1}{(a^2+b^2)^2} \tag{9.141}$$

then,

$$\frac{3(a^2+b^2)}{(a^2+b^2)^3} \quad \geq \quad \frac{4a^2-(a^2+b^2)}{(a^2+b^2)^3} \tag{9.142}$$

$$\frac{3(a^2+b^2)}{(a^2+b^2)^3} \quad \geq \quad \frac{3a^2-b^2}{(a^2+b^2)^3} \tag{9.143}$$

This also is true for a, b real.

Considering coefficients of z^{-3}, if

$$\frac{4}{(a^2+b^2)^{\frac{5}{2}}} \geq \frac{8a^3}{(a^2+b^2)^4} - \frac{4a}{(a^2+b^2)^3} \tag{9.144}$$

then,

$$\frac{4}{(a^2+b^2)^{\frac{5}{2}}} \quad \geq \quad \frac{8a^3-4a^3-4ab^2}{(a^2+b^2)^4} \tag{9.145}$$

$$\geq \quad \frac{4a(a^2-b^2)}{(a^2+b^2)^4} \tag{9.146}$$

$$\geq \quad \frac{4a}{(a^2+b^2)^3} \tag{9.147}$$

$$\frac{16(a^2+b^2)}{(a^2+b^2)^6} \quad \geq \quad \frac{16a^2}{(a^2+b^2)^6} \tag{9.148}$$

This is true for a, b real.

Considering coefficients of z^{-4}, if

$$\frac{5}{(a^2+b^2)^3} \geq \frac{32a^5}{(a^2+b^2)^6} - \frac{32a^3}{(a^2+b^2)^5} + \frac{6a}{(a^2+b^2)^4} \tag{9.149}$$

then

$$\frac{5}{(a^2+b^2)^3} \geq \frac{32a^5 - 32a^3(a^2+b^2) + 6a(a^2+b^2)^2}{(a^2+b^2)^6} \tag{9.150}$$

$$\frac{5(a^2+b^2)^3}{(a^2+b^2)^6} \geq \frac{-32a^3b^2 + 6a^5 + 6ab^4 + 12a^3b^2}{(a^2+b^2)^6} \tag{9.151}$$

$$\frac{5(a^6+3a^4b^2+3a^2b^4+b^6)}{(a^2+b^2)^6} \geq \frac{2(3a^5 - 10a^3b^2 + 3b^4)}{(a^2+b^2)^6} \tag{9.152}$$

$$\frac{a(5a^5+15a^3b^2+15ab^4+5\frac{b^6}{a})}{(a^2+b^2)^6} \geq \frac{6a^5 - 20a^3b^2 + 6b^4)}{(a^2+b^2)^6} \tag{9.153}$$

It is possible to see that this is true when it is considered that $|a|, |b| > 1$ since $5a^6$ can never be less than one less than $6a^5$, $15a^3b^2$ is much greater than unity and $-20a^3b^2$ is less than zero and $15ab^4$ is greater than $6b^4$. Unfortunately this rise in complexity is likely to continue as the coefficient number increases and hence it is unlikely to be possible to demonstrate the general case:

$$\frac{n}{(a^2+b^2)^{\frac{n}{2}}} \geq \sum_{k=1}^{K} \begin{pmatrix} n-k \\ k-1 \end{pmatrix} \frac{(-2a)^{n-2k+1}}{(a^2+b^2)^{n-k+1}}(-1)^{n-k} \tag{9.154}$$

One method that initially looks promising is to rearrange the problem to:

$$\frac{n(a^2+b^2)^{\frac{n}{2}}}{(a^2+b^2)^n} \geq \sum_{k=1}^{K} \begin{pmatrix} n-k \\ k-1 \end{pmatrix} \frac{(-2a)^{n-2k+1}(a^2+b^2)^k}{(a^2+b^2)^{n+1}}(-1)^{n-k} \tag{9.155}$$

Unfortunately it is not easy to expand $(a^2+b^2)^{\frac{n}{2}}$ since $\frac{n}{2}$ is not a natural number.

As stated before the problems encountered with this analysis may well yield to further work.

9.7 Measurements for BL Linearity Estimation

These are the results of obtained during the first pass.

Mass (g)	Location (mm)	Current (A)
0	4.75	0
10	4.65	.015
20	4.50	.037
30	4.40	.062
50	4.15	.112
70	3.90	.162
100	3.60	.219
150	3.10	.328
200	2.65	.424
250	2.25	.535
300	1.80	.672
350	1.50	.760

These are the results obtained during the second pass. Note that the location at rest is different since a different thickness plate was used. The plate used in the first run was of approximately 2mm light alloy with a large number of small holes punched in it a regular intervals, resulting in a light, rigid plate. Since a punching process was used it is possible that the plate was not flat and this may have been significant. The second pass used a significantly thinner but sold plate. As can be seen from the results there is no discernible difference between the results obtained.

Mass(g)	Location (mm)	Current (A)
0	3.85	0
10	3.75	.027
20	3.65	.046
30	3.52	.075
50	3.28	.127
70	3.10	.157
100	2.75	.230
150	2.30	.325
200	1.85	.425
250	1.37	.559
300	1.02	.662
350	0.70	.747

9.8 Matlab Re-sampler Code

```
% This is to be fed with a test signal and a velocity signal. Later
% versions to include velocity signal generation based on the signal.
%
% The signal is to be called "Ssig" and the velocity signal "Vsig"
% Vsig is to be measured in m/s. If Vsig is not defined vel and Vcont
% are used.

if exist('Vsig')==0,
   disp('Vsig not defined, using a scaled vel');
   Vsig = vel.*Vcont;
   temp = num2str(max(Vsig));
   disp(['Max value of Vsig is' temp]);
else
   disp('Using Vsig');
end

len_buf = 201; % the length of the buffer is +-(len_buf-1)/2 samples.
if (rem(len_buf,2)==0),  % Obviously len_buf must be odd.
  disp('warning - len_buf must be odd');
  return;
end
del = floor(len_buf/2);

t_rel = -(len_buf-1)/2:(len_buf-1)/2;    % This is an axis of
% relative time locations.

buf = zeros(len_buf,1); % sample buffer pre-filled with zeros
t_shift = 0; % Initially the time shift is zero

i=1;   % i is the sample number
for i = 1 : del,
```

```
  buf(2:length(buf)) = buf(1:length(buf)-1); % index sample buffer.
  buf(1,1) = Ssig(i);
end


i=1;   % i is the sample number
for i = 1 : length(Ssig)-del,
  buf(2:length(buf)) = buf(1:length(buf)-1); % index sample buffer.
  buf(1,1) = Ssig(i+del);
  t_shift = t_shift+(1-((c-Vsig(i))/c))-0.00001*sign(t_shift);

  t_shift_s = num2str(t_shift);
  index = num2str(i);
  disp(['index ' index '    t_shift ' t_shift_s]);

  s_out(i) = interp1(t_rel,buf,t_shift);
  t_axis2(i) = i-t_shift;
  %pause
end
```