Acoustic power flow into the ear
and the auditory microstructure

A thesis submitted for the degree of
Doctor of Philosophy
of the
University of London

Antony James Steven Locke
Institute of Laryngology and Otology
University College London Medical School
Abstract

An experimental technique to determine the acoustic power absorbed by the human ear at absolute threshold is described and applied to data recorded in adult subjects. A previously published method of electroacoustic probe calibration in terms of equivalent Thevenin source parameters is substantially ameliorated.

Careful and detailed measurements of continuous tonal aural sound pressure (CTASP) are presented. Ear canal input impedance, reflectance and absolute power flow constituents are derived from CTASP data.

Auditory microstructure, characterised by spectral periodicity, is observed and validated in CTASP, impedance, reflectance and power flow parameters at a 20 dB SPL stimulus level, but undetectable at 60 dB SPL. Periodicity in the ear canal acoustic parameters elicited at low stimulus levels is found to be commensurate with absolute threshold microstructure.

An elementary analogue network model of the peripheral auditory system is formulated, enabling cochlear input impedance and reflectance to be inferred from ear canal acoustic parameters. At a 20 dB SPL stimulus level a non-zero cochlear reflectance is inferred, implying that energy propagates basally, as well as, apically. Microstructure amplitude in cochlear input impedance is shown to be 4 dB greater than that in ear canal input impedance, a consequence of decoupling of the probe from the tympanic membrane.

A proportionality between transmittance and auditory sensitivity exists, implying that the ear couples more efficiently to the sound source, and consequently extracts proportionally more power, at peaks in sensitivity. However, the measured change in coupling is inadequate to wholly explain threshold microstructure. An explanation is offered by applying empirical data to a phenomenological model of power flow within the peripheral auditory system. It is argued that threshold microstructure arises predominately from a phasic interaction of the basalward and apical travelling waves effectively modifying the spatial distribution of energy within the cochlea.
Acknowledgements

I would like to thank Prof. David Kemp and Dr. David Brass for their supervision and support throughout this project. Thanks also to the experimental subjects for their time and patience.

A special thanks go to Gail, Mum and Dad and Jon for their encouragement over the past years.

Also, big thanks go to Jenny Hill, Katherine Denman-Johnson, Andrew Simpson and Siobhan Ryan for lunch-breaks.

This work was supported by the Medical Research Council. The auditory biophysics (ILO) fund provided additional financial support.
Table of contents

1.0 Introduction
   1.1 Anatomy and function of the ear 11
   1.2 Specification of the peripheral auditory system 14

2.0 Literature Review
   2.1 Ripples in the audiogram 17
   2.2 Microstructure in equal loudness data 19
   2.3 Auditory microstructure observable in the ear canal 20
   2.4 Periodicity in auditory microstructure 23
   2.5 Coincidence in frequency location of auditory microstructure extrema 25
   2.6 Physical basis for auditory microstructure 27
   2.7 The nature of wave motion in the cochlea 28
   2.8 Wave reflection model of auditory microstructure 29
   2.9 The active cochlea 31
   2.10 Modification of cochlear input impedance due to the basal travelling wave 33
   2.11 Method to measure ear canal input impedance 35
   2.12 Method to estimate Thevenin parameters 36
   2.13 Ear canal acoustic measurements 38
   2.14 Summary 39
   2.15 Scope of thesis 40
   2.16 Outline of thesis 41

3.0 Probe assembly calibration: Theory 43
   3.1 Formalism of the least mean squares solution of an overdetermined set of system equations 43
   3.2 Definition of an over-determined system 43
   3.3 Matrix representation of overdetermined system equations 44
   3.4 Principle of least mean squares 45
   3.5 Graphical construction approach to method of least mean squares 45
   3.6 Variance and Standard errors 47
   3.7 Minimisation of error function 49
4.0 Probe assembly calibration: Methods
   4.1 Probe construction
   4.2 Choice of reference loads
   4.3 Mathematical modelling of acoustic loads
   4.4 Construction of reference loads
   4.5 Stimulus generation and data acquisition
   4.6 Propagating and evanescent modes
   4.7 The effect of the evanescent wave upon acoustic measurements

5.0: Probe assembly calibration: Results
   5.1 Sound pressure measurements
   5.2 Measuring microphone frequency response
   5.3 Electroacoustic probe Thevenin parameters
   5.4 Errors in calibration
   5.5 Varying sound source constructions
   5.6 Making input impedance measurements
   5.7 Propagation of errors
   5.8 Measured input impedance of hard walled rigidly terminated tube
   5.9 Measured input impedance of Brüel & Kjær ear simulator
   5.10 Measured input impedance of adult ears

6.0: Ear canal acoustic measurements
   6.1 Introduction
   6.2 Nature and influence of Spontaneous OAEs
   6.3 Subjects
   6.4 Influence of the probe upon SOAEs
   6.5 Ear canal acoustic measurement system
   6.6 Continuous tonal aural data
   6.7 CTASP spectral periodicity
   6.8 Measurements of CTASP at different stimulus levels
   6.9 Network analysis of CTASP data
   6.10 Nature of the subtracted sound pressure
   6.11 Sources of error in the subtracted sound pressure
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.12 Summary</td>
<td>117</td>
</tr>
<tr>
<td>7.0 Derived ear canal parameters</td>
<td>118</td>
</tr>
<tr>
<td>7.1 Introduction</td>
<td>118</td>
</tr>
<tr>
<td>7.2 Definition of network parameters</td>
<td>118</td>
</tr>
<tr>
<td>7.3 Definition of scattering parameters for one port networks as applied to ear canal acoustics</td>
<td>119</td>
</tr>
<tr>
<td>7.4 Power flow in a one port network as applied to ear canal acoustics</td>
<td>120</td>
</tr>
<tr>
<td>7.5 Power flow constituents: an alternative scheme</td>
<td>121</td>
</tr>
<tr>
<td>7.6 Ear canal input impedance: empirical data</td>
<td>122</td>
</tr>
<tr>
<td>7.7 Ear canal reflectance</td>
<td>127</td>
</tr>
<tr>
<td>7.8 Power flow in the ear canal</td>
<td>129</td>
</tr>
<tr>
<td>7.9 Summary</td>
<td>131</td>
</tr>
<tr>
<td>8.0 Cochlear parameters inferred from ear canal measurements</td>
<td>132</td>
</tr>
<tr>
<td>8.1 Introduction</td>
<td>132</td>
</tr>
<tr>
<td>8.2 Phenomenological modelling approach: 2-Port network theory</td>
<td>133</td>
</tr>
<tr>
<td>8.3 Deconstructing the peripheral auditory system</td>
<td>137</td>
</tr>
<tr>
<td>8.3.1 Representation of the acoustic response of electroacoustic probe</td>
<td>138</td>
</tr>
<tr>
<td>8.3.2 Acoustic description of the occluded ear canal</td>
<td>138</td>
</tr>
<tr>
<td>8.3.3 Modelling the mechanics of the middle ear complex</td>
<td>139</td>
</tr>
<tr>
<td>8.3.4 Range of validity</td>
<td>142</td>
</tr>
<tr>
<td>8.4 Cochlear input impedance inferred from ear canal input impedance</td>
<td>144</td>
</tr>
<tr>
<td>8.5 Physical basis of the origin of microstructure</td>
<td>146</td>
</tr>
<tr>
<td>8.5.1 Theoretical treatment of cochlear input impedance</td>
<td>147</td>
</tr>
<tr>
<td>8.5.2 High level linear regime</td>
<td>148</td>
</tr>
<tr>
<td>8.5.3 Low level linear regime: Role of cochlear pressure reflectance</td>
<td>152</td>
</tr>
<tr>
<td>8.5.4 Nature of cochlear pressure reflectance</td>
<td>153</td>
</tr>
<tr>
<td>8.6 Two-port network model incorporating scattering parameters</td>
<td>156</td>
</tr>
<tr>
<td>8.7 Inferred cochlear pressure reflectance</td>
<td>159</td>
</tr>
<tr>
<td>8.8 Modelling stimulus frequency otoacoustic emissions</td>
<td>163</td>
</tr>
<tr>
<td>8.9 Utility of group delay applied to ear canal acoustic measures</td>
<td>166</td>
</tr>
<tr>
<td>8.10 Summary</td>
<td>169</td>
</tr>
</tbody>
</table>

page 6
List of figures

Figure 2.1: Analogue of an electroacoustic probe coupled to an acoustic load 36

Figure 3.1: Schematic construction illustrating the relationship between the vectors under the least squares condition 46

Figure 4.1: Input impedance of each tube in the calibration tube set. The parameter in each panel refers to the tube listed in Table 4.1 55

Figure 4.2: Network analogue illustrating the effect of the additional evanescent term 60

Figure 5.1: Schematic of the stimulus delivery and sound pressure measurement system used for Thevenin electroacoustic probe calibration 62

Figure 5.2: Difference in sampling at the probe tip compared to the terminating disc 64

Figure 5.3: Frequency response of two microphones under white noise free-field conditions 66

Figure 5.4: Thevenin parameters of the Otodynamics BS probe against frequency 68

Figure 5.5: Standardised residual errors pre-minimisation 72

Figure 5.6: Standardised residual errors post-minimisation 73

Figure 5.7: Effect of acoustic resistor upon Thevenin parameters 75

Figure 5.8: Effect of length of port tubing upon Thevenin parameters 76

Figure 5.9: Measured input impedance of a cylindrical hard walled tube 79

Figure 5.10: Measured input impedance of Brüel and Kjær ear stimulator Type 4157 81

Figure 5.11: Measured input impedance of adult ear 83

Figure 5.12: Effect of the occluded ear canal length upon the ear input impedance 85

Figure 6.1: Narrowband spectral analysis of ear canal sound pressure against frequency for each subject 90

Figure 6.2: Schematic of the experimental set-up 94

Figure 6.3: Sound pressure measured in the occluded ear canal, \( p_{cc}(\omega, A) \), of subject SS at two stimulus levels 96
Figure 6.4: Sound pressure measured in the Brüel and Kjær ear simulator Type 4157, at two stimulus levels

Figure 6.5: CTASP data for four subjects

Figure 6.6: Quantitative analysis of CTASP microstructure periodicity

Figure 6.7: CTASP data at various stimulus levels

Figure 6.8: Block diagram relating network parameters applicable for the analysis of CTASP data

Figure 6.9: Transfer function derived from CTASP data shown in figure 6.3

Figure 6.10: Vector representation, from Kemp and Chum (1980), of $p_{ec}[\omega, A]$ at two stimulus levels

Figure 6.11: The subtracted sound pressure, $p_{em}[\omega, A_{L}]$, for two female subjects

Figure 7.1: Calculated ear canal acoustic input impedance in the plane of the probe assembly tip, $Z_{ec}[\omega, A]$, at two stimulus levels

Figure 7.2: Normalised ear canal resistance and reactance against frequency at two stimulus levels

Figure 7.3: Ear canal reflectance against frequency at two stimulus levels

Figure 7.4: Ear canal power constituents, incident, reflected and absorbed, at two stimulus levels

Figure 8.1: a) Representation of a generic linear n-ported network, $^{1}N_{n}$. b) When only two pairs of independent terminals, or ports, can be identified the formalism reduces to a 2-port network, $^{1}N_{2}$

Figure 8.2: Deconstructing the peripheral auditory system

Figure 8.3: Network topology of Zwislocki analogue network of middle ear less the stapes/cochlear complex branch, relating each block on the basis of functional anatomy

Figure 8.4: Cochlear input impedance inferred from ear canal input impedance at two stimulus levels

Figure 8.5: Model cochlear and ear canal input impedance as a function of cochlear pressure reflectance

Figure 8.6: Inferred cochlear pressure reflectance from derived ear canal pressure reflectance
Figure 8.7: Simulated stimulus frequency otoacoustic emission as a function of cochlear pressure reflectance 165
Figure 8.8: Network transfer function for a second order bandstop filter 167

Figure 9.1: Example of subject’s response during an adaptive up-down procedure 174
Figure 9.2: Measurements of pure tone absolute auditory threshold microstructure 176
Figure 9.3: Iso-sensation ear canal power flow 183
Figure 9.4: Phenomenological model of power flow in the peripheral auditory system 185
Figure 9.5: Section of cochlear duct, of length $\Delta x$, modelled as a one dimensional linear transmission line 188
Figure 9.6: Illustration of standing waves within the cochlea 190

List of tables

Table 4.1: Dimensions of tubes used in calibration routine 52
Table 4.2: Thermodynamic constants used in the modelling of acoustic loads 54
Table 4.3: Table of inflection points of the $m^{th}$ order Bessel function of the first kind 58
Table 4.4: $1/e$ distances for circular tube, $a=4.5\, \text{mm}$ 59

Table 5.1: Quantitative analysis of standardised residual errors 74

Table 7.1: Thermodynamic constants appropriate for ear canal acoustics 125

Table 8.1: Parameters of the average, standardised, occluded ear canal (ANSI S3.25-1979) 139
Chapter 1: Introduction

1.1 Anatomy and function of the ear

The ear is a transducer: sound energy approaching the pinna is transformed and encoded into electrical impulses on the auditory nerve. Therefore, not only is the ear a transducer in the physical domain, but a sensory organ in a biological sense. The human ear is made up of three components, the outer, middle and inner ears. The mode of energy transfer provides a useful method for the demarcation between the three components. Anatomical and functional descriptions, with increasing levels of detail, can be found in Kinsler et al (1982), Pickles (1988) and de Boer (1980, 1984 and 1991).

A proportion of the sound impinging upon the pinna of the outer ear is absorbed, providing the acoustic input to the peripheral auditory system. Sound propagates, via the ear canal, to its oblique medial termination known as the tympanic membrane (TM), inducing motion of the TM, thereby transforming sound energy to mechanical energy.

The middle ear occupies an air filled irregular volume, the tympanic cavity (TC). The TC is fenestrated by three windows; the TM, oval window (OW) and round window (RW). The Eustachian tube couples the middle ear with the pharynx and regulates the static pressure within the tympanic cavity, such that under normal function the net static pressure across the TM is zero. The tympanic cavity is traversed by a chain of three mobile bones, known as the ossicles; the malleus whose vertical process, the manubrium, is directly coupled to the TM, the incus and the stapes. With the incus located between the malleus and stapes, the trio of ossicles provide a mechanical linkage and transmission path between the TM and OW.

---

1 Throughout this thesis, unless where specified, reference to the ear is taken to mean only those structures involved in the auditory process in man are considered. Components of the vestibular system are ignored. Therefore, the ear and the peripheral auditory system are synonymous.

2 Reference to sound is used with caution due to semantic ambiguity; does sound refer to a physical quantity or a psychological phenomenon? In this study, sound refers to 1) a wave phenomenon resulting from oscillations of pressure, particle velocity, stress etc., which 2) evokes an auditory sensation. It is noted that not all phenomena conforming to 1) lead to 2), for example, ultrasound. Additionally, sound provides a distinction between static, d.c., pressure and fluctuating pressure.
Chapter 1: Introduction

An interior view from the stapes reveals the fluid filled inner ear, and specifically the bulbous entrance chamber known as the vestibule. The inner ear, embedded deep in the temporal bone, incorporates the functional, sensory units, or organs, for both audition and the vestibular system. Located within the inner ear is the cochlea, a logarithmic spiral shaped organ. The cochlea displays a high degree of cellular spatial order, which is able to detect minute changes in pressure. Spiralling along the cochlea are three canals, termed scalae; the scala vestibuli (SV) which is continuous with the vestibule, the scala tympani (ST) and scala media (SM). The Reissner's membrane separates the SV from SM, while it is important to state for reasons of cellular electrophysiology outlined below, that the boundary between the SM and ST is the reticular lamina (RL) \(^3\) (Dallos, 1992). Running along the length of the cochlea is the basilar membrane (BM), a rugose, fibrous, corrugated membrane. Attached to one surface of the spiralling BM is the organ of Corti, a highly specialised cellular matrix which consists of sensory hair cells, numbering approximately 15,000 per cochlea and a number of differentiated supporting cells. The organ of Corti together with the BM, RL, fluid chamber in the tunnel of Corti and a gelatinous flap, known as the tectorial membrane, make up the cochlear partition (CP).

Vibration of the OW effectively launches a water wave in the cochlear fluid of the scala vestibuli, resulting in a hydromechanical wave which propagates apically. The BM exhibits a highly graded mechanical characteristic which acts in similar fashion to a nonuniform tapering transmission line. Such a system supports a travelling wave, whose speed and wavelength under harmonic stimulation decreases as the wave travels towards a point along the BM which is preferential for that frequency, known as the characteristic place. Energy at different frequencies is concentrated onto different characteristic places, a property referred to as tonotopicity; displacements of the cochlear partition at the apex are maximal during low frequency harmonic vibration at the OW, while high frequencies maximally displace the cochlear partition more towards the base. In essence, the cochlea acts as an acoustic prism.

\(^3\) It is often stated, incorrectly, that the basilar membrane separates the SM and ST (Pickles, 1988; Gulick \textit{et al}, 1989).
Chapter 1: Introduction

The sensory cells are categorised into two groups, inner hair cells (IHC) and outer hair cells (OHC), whose membranes are semipermeable, i.e., only particular ions can penetrate the membrane. In common with neurones, the intracellular potential is negative with respect to the immediate environment. On the upper surface of both the IHC and OHC is an array of narrow cylindrical structures referred to as stereocilia. Significantly, the stereocilia are bathed in an endolymphatic fluid while the base of the cell is swamped in perilymph, with the reticular lamina forming the ionic barrier between the two fluids. Endolymph and perilymph, possess distinct ionic composition and therefore have different electrical properties resulting in a net electrical potential across the membrane of the hair cells.

Displacement of the IHC stereocilia results in a change in intracellular potential. However, the magnitude of the change is maximal to radial displacements, i.e., movements perpendicular to the direction of the travelling wave. Displacement towards the lateral side, towards the stria vascularis, depolarises the cell, while movement in the opposite direction, towards the spiral limbus, leads to hyperpolarisation. This observation provides the basis of the idea that lateral displacement of the stereocilia enable ion (K⁺) flow into the cell which depolarises the cell, hence the usage of the term ‘ion channels’ to describe the mechanism. At the base of the cell are voltage dependent calcium channels which govern the release of neurotransmitters that act onto the dendrites of the afferent neurones, generating an action potential in the fibres that make up the auditory nerve.
1.2 Specification of the peripheral auditory system

Sounds that are significant to humans vary in amplitude, phase and frequency with time. For instance, a speech waveform is a continuously spectrally changing acoustic signal. From a design perspective, the ear must be able to initiate an array of neural spikes in the fibres, that make up the auditory nerve, in a spatial and temporal fashion, in order to represent the time varying stimulus. The temporal and spectral characteristics of a signal are reciprocally related. At one extreme, an impulse of infinitesimally short duration is capable of delivering energy at an infinite number of frequencies. In contrast at the other extreme, energy at a single frequency, or spectral component, requires a sinusoidal oscillation of infinitely long duration. Speech and other naturally occurring sounds lie on the continuum somewhere between these extrema.

In general, the constraints governing the analysis of signals reflect this reciprocal relationship, since high frequency resolution is achieved by sampling the signal for a long duration, with narrow filters. Conversely, higher temporal resolution is achieved with broader filters. It has been stated that the ear has evolved to find a near optimal solution to these conflicting requirements (Dallos, 1992).

Quantifying auditory temporal resolution, Plomp (1964) found that the gap threshold, the shortest duration of silence that was perceptible between two noise bursts, was typically 2-3 ms.

To describe the frequency analysis capabilities of the ear, or more properly the whole auditory system, reference is made to two auditory measures; frequency resolution and frequency discrimination.

The ability to resolve a signal in the presence of noise relates to the measure of frequency resolution, such that the subject has to filter out noise in order to detect the signal. The width, or resolution of the filter determines the resolving power and the quality of frequency analysis of the auditory system. For normal hearing humans, the 3
dB bandwidths are typically 10 to 15%, equivalent to a whole musical tone\(^4\) (Moore and Glasberg, 1983).

Frequency discrimination, or frequency difference limen, is defined as the minimum difference in frequency that is perceptible between two tones presented one after another. Under this paradigm, a frequency change of 0.2 to 0.3%, equivalent to 1/25th of a semitone, can be discriminated by normal hearing subjects.

There is a numerical disparity in the parameters yielded for each measure, reflecting two different paradigms in the presentation of the stimuli. Frequency resolution reflects the spectral capabilities only of the auditory system and is a measure which is used in many physical fields, whereas, the frequency difference limen is a measure of the ability to discriminate between two neural patterns at different times.

The range between the lowest and highest frequency of pure tones which elicit a sensation of pitch defines the auditory frequency range. The frequency range of the human ear from 0.02 to 20 kHz\(^5\) is relatively considerable, spanning 10 octaves. In contrast, the eye operates over only 1 octave and the olfactory system approximately 3.5 octaves (Turin, 1997). Encoding of acoustic waves with air-borne wavelengths between 17.5 m to 1.75 cm takes place in the human cochlea, which has an uncoiled length of 3 cm. Therefore, sound waves that are perceptible to humans have wavelengths comparable to dimensions of everyday objects in the environment, which is significant due to diffraction effects.

The lowest stimulus level at which there is an auditory sensation is described by the threshold of hearing. The greatest sensitivity, that is the lowest threshold of hearing, is

\(^4\) For the even tempered musical scale the ratio of two frequencies separated by a semitone is defined as \(2^{1/12}\). This equates to a percent change of 5.9%.

\(^5\) From studies by Békésy (1960), it was demonstrated that sinusoidal stimuli of frequencies below 20 Hz were perceptible. However, at 5 Hz threshold was roughly 120 dB SPL, leading to the possibility of the generation of aural harmonics which may cue perception. Additionally, it was observed that at such frequencies and intensities, the stimulus did not elicit a pitch, but rather a percept akin to a tactile sensation.
in the mid, 1-2 kHz, frequency range\(^6\), corresponding to a power incident upon the TM of 50 aW. At the threshold of hearing, the corresponding displacement amplitudes of the structures in the auditory periphery are subatomic, fractions of an ångström \(10^{-10}\) m. Thermal noise arising from the displacements of the cilia of hair cells is estimated to be of the same order of magnitude, leading to the conclusion that the sensitivity of the auditory system is not limited by a design compromise, but by the physical limits imposed by Nature.

The auditory system exhibits a relatively large dynamic range, defined as, the ratio of the stimulus level at which the extraction of useful information ceases\(^7\), to the threshold of hearing. The power incident upon the TM corresponding to the upper limit is approximately 50 \(\mu\)W, that is, of the order of one million million times greater (120 dB) than the minimum power required to elicit an auditory sensation.

\(^6\) As a note of interest, the 2 kHz octave band corresponds to the frequency range that conveys the most information in speech.

\(^7\) Implied in the use of the term 'extraction of useful information' is the notion that sounds impinging upon the ear are ultimately analysed in terms of the information which is carried in the signal.
Chapter 2: Literature review

2.1 Ripples in the audiogram

Elliot (1958) described a ripple effect in the pure tone audiogram of normal unimpaired human subjects. Absolute thresholds of the right ears of two subjects were presented, although the experimental methodology was not described except that the audiogram was measured from 0.4 to 3 kHz in 10 Hz steps. The resulting audiograms showed a regular and repeatable characteristic pattern of undulations with frequency, exhibiting threshold variations of up to 10 dB. Elliot reported that the ripple of the left ear differed from that of the right, implying that the pattern is unique to the ear under test. Additionally, the ripple pattern in the audiogram of one subject did not match, either in frequency location of extrema or ratio of maxima to minima, the audiogram of a second subject. The ripple pattern present in the audiogram shows no interaural or intersubject correlation. Elliot stated that the pattern of extrema was stable, reproducible and repeatable over a time course of at least two months. Due to the repeatability, the ripple must be interpreted as a signal reflecting the response of a physical, physiological or psychological mechanism, rather than noise, or experimental error, of an uncorrelated random characteristic.

Van den Brink (1970), Thomas (1975) and Kemp and Martin (1976) report similar findings of an undulation in the pure tone audiogram. Van den Brink (1970), while principally investigating binaural diplacusis\(^1\), found large fluctuations in the audiogram for small changes in signal frequency. Thomas (1975) states that the difference in intensity between adjacent peaks and troughs is typically 12 dB. Additionally, the pitch of the stimulus remained constant as the frequency was swept between peaks, but jumped in a quantal fashion when the frequency was swept across a peak. Kemp and Martin (1976) predicted an active mechanism in order to account for the high Q-factors of the ripples in the audiogram.

---

\(^1\) An auditory phenomenon in which a fixed frequency stimulus presented binaurally elicits different pitches between left and right ears.
Wilson (1980b) observed that threshold ripple was dependent upon intracochlear static pressure, a factor which affects the impedance presented at the stapes footplate. Changes in hydrostatic pressure, assumed to induce changes in intracochlear static pressure, were achieved by tilting the body to another posture. When the body was inverted the position of the peaks and troughs reversed, while tilting the body at one particular angle obliterated the ripple.

Long (1984) studied both absolute and masked thresholds and demonstrated that the ripple flattens out with increasing masker level. For simultaneous masking with broadband noise no undulation was seen when masked thresholds were above 40 to 50 dB SPL re 20 \( \mu \text{Pa} \). Additionally, it was shown that the pattern of undulation in the threshold was invariable over a period of 18 months.

Cohen (1982) implemented a rigid, fixed choice psychoacoustic paradigm in a detailed study of auditory threshold across a narrow frequency sweep. However, it was concluded that neither the frequency spacing between neighbouring maxima and minima, nor the ratio of maxima to minima exhibited an observable regular pattern. Such a finding suggests that the auditory threshold ripple is not universal, but rather is present in a proportion of subjects.

Here, auditory microstructure is defined as a characteristic of a hearing phenomenon which exhibits a reproducible locally quasiperiodic series of maxima and minima with frequency\(^3\), manifesting a spectral periodicity equal to the frequency taken to complete one cycle, and spectral amplitude determined as the ratio of maximum to minimum within one cycle. The ripple in the audiogram is therefore an example of auditory microstructure.

---

\(^2\) From now on we define dB SPL as the sound pressure level in decibels referenced to 20 pPa rms.

\(^3\) A note on nomenclature. Fine structure, rippling effect, undulation and irregularities are, within this report, taken to be synonymous with microstructure.
2.2 Microstructure in equal loudness data

Kemp (1979a) conducted a detailed study of auditory microstructure. As well as pure tone absolute threshold, suprathreshold psychoacoustic measurements such as, equal loudness experiments and frequency discrimination tasks were undertaken. During equal loudness experiments, the subject was instructed to vary the level of a tone at one frequency so that it matched the loudness of a fixed level tone at another, reference frequency. The procedure was repeated for tones of frequencies around the reference tone. The resulting curve of the subject’s level responses give an equal loudness contour. At low levels, i.e., reference levels of 5 dB SPL, the equal loudness contour exhibited pronounced microstructure. As the level of the reference tone was increased, in steps of 5 dB SPL, the ratio of maxima to minima amplitude decreased. At 40 dB SPL the equal loudness contour exhibited no undulation.

Due to the lack of proportionality between the amplitude of microstructure and applied sound pressure level, the data clearly demonstrates a compressive nonlinear auditory

4 Discussion on the definition of linearity, when applied to systems, is facilitated with reference to a Volterra series (Simon and Tomlinson, 1984), where the output, \( y(t) \), is related to the input, \( x(t) \), by a Volterra series,

\[
y(t) = h_0 + \int h_1(\tau)x(t-\tau)\,d\tau + \int \int h_2(\tau_1,\tau_2)x(t-\tau_1)x(t-\tau_2)\,d\tau_1\,d\tau_2 + \ldots
\]

where \( h_0, h_1, h_2 \ldots \) are the Volterra kernels. A linear system is defined as a special case, in which the Volterra series is truncated after the first-order kernel, \( h_1 \). The system transfer function, \( H(o) \), is then related to the system unit impulse response, \( h(t) \), by a single Fourier transform operation. Inclusion of the higher order terms of the series reflect nonlinearity in the system response. Underlying the principle of linearity is that of superposition; for a given linear system, if the output to an input \( x_1(t) \) is \( y_1(t) \) and the output to \( x_2(t) \) is \( y_2(t) \), then the output to \( a_1x_1(t) + a_2x_2(t) \) is \( a_1y_1(t) + a_2y_2(t) \) where \( a_1 \) and \( a_2 \) are arbitrary constants (Ruston and Bordogna, 1966). That is, the output to two inputs presented simultaneously, is simply the algebraic sum of the outputs when presented separately. Such a statement is practical since it states that a linear system can equally be characterised by analysing its response to either sinusoids or impulses, the two simplest but most distinct signals. However, there exists a plethora of nomenclature applied to the definition of linearity, which may be less rigorous as the definition above, but are more applicable to when the input is sinusoidal. Further definitions separate the previous definition into one of (restricted) superposition and homogeneity, often referred to as proportionality (Kuo, 1962). The restricted superposition drops the reference to the two arbitrary constants \( a_1 \) and \( a_2 \), and is more properly referred to as additivity. The level dependence is then governed by the principle of homogeneity; if the input level is changed, by a factor \( k \), which gives rise to a corresponding change, \( k \), in the output level, then homogeneity is preserved. Often the principle of additivity, rather than applied to signals of arbitrary complexity, is restricted to two sinusoidal inputs of different frequencies (Moore, 1986); such that, the response of a system to two frequencies presented simultaneously is equal to the sum of the responses when the frequencies are presented separately. For mechanical systems, linearity implies that vibration is harmonic if the return force is directly proportional to displacement (Main, 1994). If the dependency is different then the system is nonlinear and supports vibrations which are
phenomenon which is most prominent at low stimulus levels and is obliterated at moderate to high levels.

The data discussed above support the view conjectured by Kemp (1979a) that auditory parameters such as pure tone threshold, loudness and pitch are not smooth, slowly varying functions of frequency, but display a microstructure which, in the case of absolute thresholds, is not observed in conventional audiometry.

The microstructure observable in the pure tone audiogram, masked thresholds and equal loudness data required experimental paradigms that are psychoacoustic, i.e., required a subjective action in response to acoustical stimulation. This leads to the question, "is microstructure observable in a purely physical measure of an auditory phenomenon?".

2.3 Auditory microstructure observable in the ear canal

Microstructure is observable in low level sweeps across frequency of raw ear canal sound pressure measurements, (Kemp, 1979a; Kemp and Chum, 1980; Wilson, 1980ab; Zwicker and Schloth, 1984). The experimental method reported in Zwicker and Schloth (1984) required measurement of the vector sound pressure in the ear canal during continuous tonal stimulation from a high impedance source under constant electrical drive conditions. The stimulus was swept slowly in frequency, between 1350 and 1550 Hz, and at various stimulus levels; 0 to 35 dB SL \(^5\) in 5 dB steps. The greatest fluctuation in both magnitude and phase can be seen at 0 dB SL with the spectral ripple abolished at a stimulus level of 60 dB SL.

It was observed that the phase curves at the lowest stimulus levels are identical, as well as an invariance in the shape of the magnitude curves, but which are separated by an interval corresponding to the stimulus level increment, a homogeneity which befits a

\(^{anharmonic. However, the response of the transverse BM velocity or displacement to harmonic stimulation is not ideally linear but resembles to a certain degree, as if it exhibited, a linear characteristic. Hence the use of the term quasilinear, defined as the response to a sinusoid is largely sinusoidal, when considering cochlear mechanics (Kanis, 1995; Brass and Kemp, 1993).}

\(^5\) dB SL is defined as the intensity in decibels above threshold of audibility. With reference to section 2.1 and specifically discussion on auditory microstructure, sensation level appears to be a rather approximate and inappropriate measure due to the often 12 dB difference between threshold maxima to minima.
linear system. The data at high stimulus levels indicate, due to the deviation from homogeneity, the presence of a nonlinearity similar to that shown in equal loudness data (Kemp, 1979a).

Such sound pressure measurements are occasionally referred to as stimulus frequency otoacoustic emissions (SFOAEs) (Zwicker, 1979 and 1990). This interpretation can appear confusing since SFOAEs usually describe a signal that is purely of cochlear origin (Brass and Kemp, 1991).

Continuous tonal sound pressures consist of two signals; one signal that reflects the gross linear response of the auditory periphery, and a second cochlear originating signal that is level dependent. A method to extract the cochlear originating signal was proposed by Kemp (1979a) by using a complex pressure, i.e., vector, subtraction method, which was developed in Kemp and Chum (1980) and utilised in Zwicker and Schloth (1984) and Guinan (1990). A pressure response at 70 dB SPL, which is sufficiently high so that the additional level dependent signal is negligible, is, with suitable scaling, subtracted from the low level response. The subtracted, or residual, signal was termed the SFOAE (Kemp and Chum, 1980).

An alternative method to extract the SFOAE by simultaneously generating a continuous stimulus tone and a pulsed probe tone of the same frequency was described in Kemp and Souter (1988) and developed in Brass and Kemp (1991). By applying the probe tone at appropriate positions in the cycle of the stimulus tone a cancellation of both the probe and stimulus was achieved to reveal a nonlinear residual at the stimulus frequency. The residual is not strictly the SFOAE, but rather reflects the suppression by the probe tone of the stimulus tone evoked SFOAE. The, nontrivial, relationship between the residual and SFOAE is given in Brass and Kemp (1991). The residual exhibits a compressive nonlinearity indicative of an OAE (Harris and Probst, 1990).

---

6 Otoacoustic emissions (OAEs) are ear canal acoustic manifestations of cochlear biomechanics. OAEs are conventionally classified in terms of the stimulus that elicits the response. Under no stimulation the category of spontaneous OAEs (SOAEs) is used. When a stimulus is applied the OAE is evoked, hence EOAE. EOAEs are further distinguished by the stimulus; distortion product (DPOAE), transient, or click, (TEOAEs) and stimulus frequency (SFOAEs).

7 Such a response recorded in the ear canal is a function of the mechanics of the ear canal, middle ear and cochlea.
Caution in the comparison of SFOAEs derived from the two paradigms, vector subtraction and probe stimulus cancellation, must be exercised. SFOAEs derived from the vector subtraction method are not exactly equivalent to SFOAEs detected under the probe stimulus cancellation paradigm. This is because the former carries information of the nonlinearity between the low to high stimuli levels, often separated by 50 dB, whereas the latter method elicits SFOAEs under stimuli that are of comparable intensities.

Due to the 'fuzziness' of terminology, we refer to the raw ear canal pressure measurements as continuous tonal aural sound pressures (CTASP), whereas reference to SFOAEs will, without clarification, imply a subtraction or cancellation procedure.

A rippling pattern in CTASP across frequency was present at levels down to -40 dB SPL (Wilson, 1980a). The existence of such an interference pattern at subthreshold stimuli levels, demonstrates that the spectral ripple phenomenon is independent of processes in, or a function of, the neural pathway, such as reflex loops, since the response is present without any nerve spike generation.

When two sinusoidal stimuli of different frequencies are applied to the ear simultaneously a number of spectral components recorded in the ear canal at frequencies not present in, but algebraically related to, the stimuli are present. The existence of such additional spectral components, referred to as distortion products (harmonic, difference and combinational tones), implies that the ear is a nonlinear forced vibrator.

Microstructure is also seen in measurements of distortion product otoacoustic emissions (DPOAE) (He and Schmiedt, 1993). In order to elicit a DPOAE, two tones, referred to as primaries, of frequency $f_1$ and $f_2$ ($f_2$ is usually taken to be greater than $f_1$), and level $l_1$ and $l_2$, constitute the stimulus presented in the ear canal. He and Schmiedt (1993), recorded the lower third order, cubic, distortion product, corresponding to a signal at a frequency of $2f_2 - f_1$. Primary levels were equal and varied from 45 to 65 dB SPL in 2.5 dB steps. Primary frequencies were changed in steps of 1/32 octave, while keeping the $f_2/f_1$ ratio constant at 1.2. Peak to valley amplitude ratios of up to 20 dB SPL are shown.
However, the peakiness, or degree, of microstructure in the DPOAE data does not manifest a level dependence with a diminishing peak to valley amplitude ratio with increasing primary level.

2.4 Periodicity in auditory microstructure

When displayed on a logarithmic frequency scale, microstructure observed in auditory measurements show neighbouring maxima and minima that are, roughly, equally spaced (Zwicker, 1990; Zweig and Shera, 1995; He and Schmiedt, 1993). Such a qualitative observation evinces an approximate spectral periodicity\(^8\) in the microstructure.

Quantitatively, oscillatory components that are constituent parts of the undulating microstructure can be extracted either by plotting a simple histogram of binned values of the interpeak spacings resulting in a probability distribution, or by performing a Fourier transform. An underlying single periodicity in the microstructure will be manifest as a clearly defined peak in the histogram or amplitude plot of the Fourier transform.

He and Schmiedt (1993), plotted a probability distribution of the interpeak frequencies of DPOAE microstructure, categorised in bins of width 1/32\(^{st}\) of an octave, which was maximal at a bin corresponding to 3/32\(^{nd}\) of an octave.

Zweig and Shera (1995) performed a Fourier transform on CTASP data and report a distribution centred on a pronounced peak equal to a frequency spacing, \(\Delta f/f\), of 1/15, equivalent to 1/10\(^{th}\) of an octave.

Schloth (1983) investigated the frequency spacing between spontaneous otoacoustic emissions (SOAE), i.e., nonevoked OAEs. The spectral analysis between 775 Hz and

\(\text{\footnotesize \(^8\) In the following discussion, spectral periodicity refers to a spectrum which when shifted along the abscissa a frequency (which is relative, rather than absolute, in this context) equal to a multiple integer of the fundamental frequency, i.e., the reciprocal of the period, superimposes onto the original spectrum. In other words, the autocorrelation function of the spectrum has peaks at multiple integers of the fundamental frequency only. The consequence of shifting the spectrum along the frequency axis a small whole number of spectral periods resulting in approximately the original spectrum, reveals a 'discrete spectral symmetry' (Zweig and Shera, 1995).}
3990 Hz, of sound pressure recorded in the occluded ear canal manifested a series of narrowband peaks, referred to as SOAEs. The frequency separation between neighbouring SOAEs, i.e., interpeak frequency, when displayed as a probability density against linear frequency histogram showed a bi-modal distribution between 70 Hz to 130 Hz.

In the language of critical-band rate, commonly referred to as the z-scale (Zwicker and Terhardt, 1980), the periodicity in CTASP was found to be equivalent to approximately 0.4 Bark (Zwicker, 1990), with an interquartile range from 0.35 to 0.46 Bark. Interpeak frequency of SOAEs when binned and represented as a probability density against Bark frequency histogram revealed a distinct peak located at 0.4 Bark (Dallmayr, 1985; Zwicker, 1989). Zweig and Shera (1995) produced a histogram of counts using SOAE data from Talmadge et al (1992) which exhibited a similar characteristic value. Around 500 Hz, the interpeak distance is roughly 40 Hz, while at 3 kHz it is approximately 250 Hz. Zwicker (1990) found that the distribution of the interpeak spacing between the threshold microstructure minima of 8 subjects peaked at 0.4 Bark.

The Bark is an auditory unit derived from psychoacoustical experiments on the tuning characteristics of the auditory system. Masking paradigms are undertaken to determine the width of the auditory filter at a particular centre frequency. The width of the filter, termed the critical bandwidth, is equivalent to one Bark. The tonotopic nature of the response of the cochlea, in which frequency is mapped spatially, warps the frequency, to a first approximation, in a logarithmic manner, with filter bandwidth approximately equal to one third of an octave. The Bark relates to a constant length along the BM corresponding to 1.1 mm, therefore 0.4 Bark corresponds to length of 0.44 mm along the BM. Plotting psychoacoustical data, such as threshold values derived from masking experiments, against the critical-band rate results in filter shapes which are invariant irrespective of frequency, therefore normalising the frequency warping effect of the auditory periphery.

The Bark unit is, itself, derived from psychoacoustical experiments, using moderate, around 50 dB SPL, stimulus levels. The critical bandwidth has been shown to be a function of frequency, with narrower filters, implying improved frequency selectivity,
for lower stimulus levels. The use of a psychoacoustic unit such as the Bark, derived at moderate stimulus levels, to quantify an attribute of a purely physical measure recorded at stimulus levels close to threshold appears to be inappropriate. For this reason reference to the Bark will be limited.

The similarity in periodicity observed in auditory thresholds, suprathreshold psychoacoustic measurements, CTASP sweeps, DPOAEs and the regular frequency separation between consecutive peaks of SOAEs suggest a common fundamental mechanism. Such a mechanism must exist in the peripheral auditory system since OAEs are preneural signals (Wilson, 1980a). The results reported above evince that auditory microstructure is a physical phenomenon observable in physical and psychoacoustical measurements which is intrinsic to, and the by-product of, the normal hearing process in the auditory periphery.

However, the regularity of the microstructure, as noted thus far is not universal, but rather is present in only a proportion of subjects (Wilson, 1980b). Shera and Zweig (1993) report data derived from CTASP measurements swept over frequency where an irregular, anomalous region is sandwiched between a regular, ordered microstructure.

2.5 Coincidence in frequency location of auditory microstructure extrema

Data from Zwicker and Schloth (1984) display that frequencies at which both the CTASP magnitude and SOAE are maximal coincide with the minima in threshold microstructure, leading to the observation that at frequencies where the ear emits most powerfully hearing acuity is greatest (Shera and Zweig, 1993).

The threshold data shown in Zwicker and Scholth (1984) is the rms electrical voltage level in decibels applied to the earphone. Alternatively, threshold data can be acquired by measuring the sound pressure level in the ear canal during stimulation at a level corresponding to threshold. Differences in the methodology of threshold measurement may give rise to differences in the derived threshold microstructure.
As Kim (1983) notes, the acoustic signal in the ear canal can be made up of three components; firstly, the component resulting from electrical stimulation of the earphone, referred to as the applied stimulus; secondly, the stimulus frequency OAE signal component; and thirdly, the spontaneous OAE component. The second and third components are of cochlear origin, as well as dependent upon the transmission characteristics of the middle ear. The applied stimulus component is common to both methods, electrical and SPL, of threshold measurement, whereas the SPL derived threshold measure carries information about the two additional cochlear originating signals.

Wilson (1980) states that, for three subjects, there was an intrasubject correlation, but no intersubject correlation found between the frequency position of the CTASP maxima and threshold minima. The maxima in the sound pressure level measured in the ear of one subject coincided with the frequency of the threshold minima, the second subject’s SPL data displayed a correspondence with threshold maxima, while the microstructure recorded from the third subject was somewhere between the two. Such a finding is, according to Wilson, not surprising since the phase relationship between the stimulus and reëmitted signal in the ear canal may be quite different at the level of the sensory mechanism.

To investigate the possible relationship between the applied stimulus, reëmitted signal and threshold microstructure, reference to the role of the sound delivery system and influence of energy reflectance has to be made.

Absolute sound pressures measured across frequency in ear canals during continuous tonal stimulation are a function of the source impedance of the sound delivery system. For a low source impedance relative to the load impedance, the sound delivery system approximates to a constant sound pressure source; sound pressure across the acoustic load is independent of the magnitude of the load impedance. Conversely, a constant volume velocity source approximation is valid when the source impedance is very much greater than the load impedance. CTASP derived data carries information relating to the sound source. This observation illustrates the proportionality between sound pressure level and load impedance magnitude for a constant volume velocity drive...
condition, leading to the conclusion that fluctuations in the microstructure of low level CTASP are maximal when recorded using a high impedance source.

Zwicker (1990) investigated the acoustic characteristics of the sound delivery system in relation to CTASP and auditory thresholds. Two probes, designated A and B, were developed, with differing source impedances. Probe A was constructed with a relatively high degree of acoustic damping, whereas probe B exhibited a lightly damped Helmholtz resonator characteristic with a resonance at 1250 Hz. Measurements were taken over a frequency range of 900 to 1100 Hz, i.e., in the compliance controlled frequency region of the source impedance of probe B. Microstructure in CTASP data was evident with both probe A and probe B, with the frequency spacing between neighbouring maxima similar for both probes. However, the data displays a shift along the frequency axis such that maxima measured with probe B coincided with minima measured with probe A.

Allen et al (1995) observed a small shift in frequency, of 8 Hz, between the frequency position of a SOAE, recorded with an in-the-ear probe assembly, and the maxima in the threshold microstructure, where the stimulus was delivered via headphones. The frequency shift was explained by the use of different transducers in each measure.

2.6 Physical basis for auditory microstructure

Auditory microstructure mirrors, to a first approximation, that of the pressure standing wave pattern seen in transmission line\textsuperscript{9} models, such as an organ pipe. Implicit in the analysis of such systems is the existence of backward, as well as forward, travelling waves (TW), the consequence of which is twofold. Firstly, power flows bidirectionally along the length of the transmission line. Secondly, the input impedance of the transmission line is modified by the presence of the backward TW.

\textsuperscript{9} A transmission line is defined as a system that supports wave motion such that propagation is constrained in a spatial sense. Note that the resulting wave motion is not necessarily one dimensional. Additionally, evanescent as well as propagating modes may arise.
Chapter 2: Literature review

The concept that the cochlea can support waves that propagate basally, as well as apically, was considered to be unrealistic and erroneous during the nascent era of measurements of basilar membrane dynamics, and as indicated by von Békésy (1960) was somewhat paradoxical. An unconventional method of cochlea stimulation can be achieved by applying the stimulus at the apex, in the proximity of the helicotrema, or by bone conducted stimulation. Measurements indicated that the TW was initiated at the conventional place where the stiffness was greatest, i.e., at the base, resulting in the uncomfortable observation that the TW propagated towards the source. This apparent preference for uni-directional wave motion can, however, be explained by noting the possible wave motions possible within the cochlea.

2.7 The nature of wave motion in the cochlea

Vibration of the stapes generates acoustical waves in the fluid filled columns of the cochlea. Peterson and Bogert (1950) and Lighthill (1991) note that, wave motion in a system of two tubes, such as that seen in the cochlea, where the very high compliant Reissener's membrane can be considered mechanically transparent, consists of two types of waves; fast and slow waves. The fast wave, as the term suggest has an extremely high propagation speed, and gives rise to equal pressure in both perilymphatic ducts resulting in zero net pressure across the cochlear partition along the length of the cochlear duct. The effect of the fast wave, or in other words the longitudinal wave, is that there is no net force upon, and therefore no motion of, the cochlear partition. In contrast, the slow wave gives rise to equal and opposite pressures across the cochlear partition, developing a pressure gradient close to the base. Due to the tapering nature of the cochlear partition mechanics the slow wave then propagates apically resulting in a hydro-mechanical wave. For modelling purposes, it is valid, especially at the basal end, that the slow wave can be approximated to a shallow wave, or long-wave approximation, since the wavelength of the TW is large compared to the height of the scalae. The result of such of an approximation is that the pressure is uniform across each traverse section of the cochlea, leading to a geometry that is for modelling purposes one-dimensional.
The energy of the hydro-mechanical wave is shared between the mass of the fluid columns, which controls the kinetic energy, and the stiffness of the cochlear partition which determines the potential energy. The basilar membrane has a unique mechanical characteristic. Firstly, the stiffness of the basilar membrane is highly graded; a four to five orders of magnitude difference between base to apex. Secondly, the basilar membrane is highly anisotropic, leading to the observation that along its axis, neighbouring sections are very weakly linked and can be considered to vibrate independently. The result is that the speed of the hydro-mechanical wave slows down, the wavelength decreases and energy ‘piles-up’ at a point which is preferential and characteristic for that frequency.

The TW apparently propagating towards the source during non-stapes stimulation of the cochlea can be explained as a result of the existence of a fast wave at the site of stimulation which travels towards the base and due to the impedance mismatch at the boundary, namely the OW and RW, leads to a conventional apically travelling hydro-mechanical wave.

2.8 Wave reflection model of auditory microstructure

Kemp (1978) postulated that reflected EOAEs are a consequence of a basal TW which is initiated by reflection of an apical TW due to micromechanical impedance discontinuities in the cochlear partition. Reflected OAEs can then be thought of as a measure of the mechanical response of cochlear structures, hence the term evoked cochlear mechanical response (ECMR) prevalent in incipient studies (Kemp, 1979b). Most generally, such a postulation provides the basis for the wave reflection model. The existence of anomalies in the mechanics of the cochlea can arise on account of anatomical disorder in the structures that make up the cochlear partition (de Boer, 1984). Cochlear models that exploit anatomical structural disorder are referred to as place-fixed models. Furthermore, Kemp (1978) hypothesised that the mechanical response of the cochlear partition, primarily in the region of the TW peak manifests a nonlinear characteristic. It is this nonlinear mechanical response that creates the micromechanical impedance discontinuity. Wave reflection models, in which the TW
itself governs the mechanism that modifies cochlear mechanics initiating a basal TW, are termed wave-fixed models.

Anatomical differences across species support this view. Primate cochleae show relatively more microstructural disorder compared to the regular ‘almost crystalline’ arrangement in the organ of Corti in rodents (Wright, 1984). The magnitude of TEOAEs in primates are correspondingly greater than in rodents. Schmiedt et al (1981) found that, even with a stimulus amplitude of 80 dB SPL peak, TEOAEs in gerbils proved difficult to record. Wit and Ritsma (1980) could not elicit TEOAEs in gerbil nor guinea pig, however, Zwicker and Manley (1981) measured reliable signals in guinea pigs.

Reverse transmission factors, in particular the role of the middle ear, may explain the difference across species. This seems unlikely due to large DPOAEs observed in rodent ears in contrast to small emissions in humans (Schmiedt and Adams 1981; Kemp 1979); the amplitude of the 2fl-f2 distortion product in humans is typically 50 to 70 dB less than the primary amplitudes (Lonsbury-Martin et al, 1990), whereas 30 to 50 dB lower in chinchillas (Zurek et al, 1982). It is noteworthy that this observation suggests a difference in the mechanisms involved in the generation of DPOAEs verses TEOAEs.

Focusing on the spectral periodicity evinced in auditory microstructure, suggests a regularity in the construction of the cochlear partition. Such an observation leads to ‘the paradox of periodicity’ (Shera and Zweig, 1993), in which chaos in the structure of the cochlear partition, revealed from anatomical studies, is at odds to the order seen in the microstructure of numerous auditory phenomena.

The wave reflection model makes explicit the notion that energy propagation along the cochlea can be modelled by a transmission line (de Boer, 1984), the analysis of which is analogous to electrical transmission line theory (Russell and Hjerpe, 1993). Under the long wave approximation, a short section of the cochlear can be modelled as an electrical network consisting of a series impedance and shunt admittance. The series

---

10 Admittance is the reciprocal of impedance. Immittance, a term coined by Bode, refers to both impedance and admittance.
impedance represents the inductance per unit length, or for the cochlea, the mass of fluid per unit length. The shunt admittance then represents the specific acoustic admittance per unit length of the cochlear partition. The basis of the local irregularity can now be made explicit as a local change in the shunt admittance compared to the gross, average change arising due to the graded mechanical property.

The effect of such a local immittance irregularity is the generation of a reflected, as well as transmitted wave. Reflected, or basal travelling waves, would give rise to a proportion of the incident energy propagating toward the OW. At the OW-TC interface, an impedance mismatch would in turn reflect a fraction of the wave back to an apically TW. However, a fraction of the basal TW incident upon the OW will, due to coupling by the middle ear, be transmitted to the ear canal, as an additional wave of cochlear origin. In other words, stimulation of healthy ears results in the re-emission of sound.

Reflection of the hydro-mechanical wave at a point along the BM is valid. However, it was argued with reference to cochlear models, that the resulting basal TW encountered reflection over and above that for the apically TW (de Boer and Viergever, 1984). The consequence of such a mechanism leads to reflection of cochlear waves which is manifestly asymmetric. As Shera and Zweig (1991) show, the existence of such an asymmetry is an artefact of the boundary conditions in the model.

2.9 The active cochlea

Wave reflection in the cochlea, with the existence of apical and basal travelling waves, leading to the phenomenon of wave interference only partially satisfies the demands that auditory microstructure data imply. Most notably is the level dependent, nonlinear\(^{11}\) ratio of maxima to minima in the undulating microstructure, which is contemporarily explained with reference to an active cochlea\(^{12}\).

---

\(^{11}\) The existence of a nonlinearity does not imply an active process. As Brass and Kemp (1993) note, such nonlinearities could reflect a passive attenuating system operating at high stimulus levels.

\(^{12}\) We use active to describe a process in which the power flowing out of a section is greater than the power flowing into the section. More specifically, there is greater power flowing past the characteristic place on the cochlear partition than flowing into the ear canal (Brass and Kemp, 1993).
It has been suggested that the existence of spontaneous OAEs, i.e., signals that are present in the ear canal without external stimulation, evince an active process in the auditory system, specifically the cochlea. However, ear canal signals could equally originate from narrowband filtered noise, due to brownian motion within the cochlea, as suggested by Allen (1980). Statistical analysis of SOAEs suggest that such signals are generated by an active nonlinear system (Bialek and Wit, 1984), however not all classes of passive nonlinear systems were postulated. Perplexingly, a number of nonmammalian vertebrates, for example the barn owl, manifest no evidence of cochlear activity, but exhibit SOAEs (Manley and Taschenberger, 1993).

Gold (1948) first proposed the existence of activity in the cochlear response to explain the high degree of frequency selectivity, or bandpass filtering, achieved in the auditory periphery. An active process manifest as a synchronous, cycle-by-cycle boost of mechanical vibration of a frequency-specific nature, acting as a high-Q resonator, is referred to generically as the cochlear amplifier (Davis, 1983). Such a biomechanical amplification process requires metabolic energy and displays a preferential operating level which is greatest at low intensities.

Brass and Kemp (1993) argue that analysis of BM transverse velocity measurements using the Mössbauer technique indicate mechanical activity in the cochlear. They compared BM transverse velocity amplitude measurements to amplitude data derived from a lossless passive model determined from measured BM transverse velocity phase data. The consequence of the presence of a cochlear amplifier, of gain of up to 40 dB, compared to the absence of such a mechanism, is a widening of dynamic range, an increase in sensitivity and a sharpening of frequency selectivity.

The compressive nonlinear nature in the microstructure is then explained as resulting from the cochlear amplifier whose relative response is maximal for low intensities. Specifically, the ratio of BM motion to stapes motion is greater for lower stimulation levels, resulting in a more pronounced undulation in the microstructure.
The second consequence of the presence of a basal TW is the modification of cochlear input impedance, from that response which corresponds only to apical wave propagation.

2.10 Modification of cochlear input impedance due to the basal travelling wave

Acoustic impedance, a measure which characterises the response of a system at a surface, is defined in acoustical terms as the complex ratio of sound pressure averaged over the surface to volume velocity through the surface.

Coupling of the ear canal to the cochlea is facilitated by the middle ear, which essentially acts as an impedance transformer. The upper limit of middle ear linearity is governed by a myogenic mechanism as well as the mechanical response of the ossicular chain and associated ligaments and muscles. In humans the myogenic mechanism, the middle ear reflex, is initiated when the stimulus intensity roughly corresponds to 80 dB SL (Gulick et al, 1989). Under high stimulus intensities (~110 dB SPL) applied to the TM, the motion of the ossicles, especially the malleus and to a lesser degree the incus, is asymmetric, concluding that a major source of nonlinearity is the malleo-incudal joint (Dahmann, 1929, as reviewed in Stevens and Davis, 1983). Møller (1974) observed asymmetry, and therefore nonlinearity, in the response of the middle ear, in particular in the motion of the stapes, at very high stimulus intensities, while using human cadavers. Rubinstein et al (1966) found that, for stimulus intensities <90 dB SPL, stapes displacement varied linearly with sound pressure level at the TM. Békésy (1960) observed a change in the vibratory pattern of the ossicular chain by increasing the stimulus intensity from below to above the threshold of feeling. Below the threshold of feeling the motion of the stapes footplate is reminiscent of a trap door, being a combination of a rotational and longitudinal movement. Above the threshold of feeling the stapes footplate vibrates in three modes, two rotational and one longitudinal. Intensities that are required to elicit a tactile percept are extremely high. Therefore, it is observed that, in the cat, the activation of the middle ear reflex determines the upper limit of linearity of the middle ear response (Guinan and Peake, 1967; Buunen and Vlaming, 1981).
The frequency response of the middle ear can be considered slowly varying in comparison to the periodicity of the microstructure. Changes in cochlear input impedance will, due to middle ear coupling, lead to changes in ear canal input impedance. This suggests that an SFOAE can be thought of as a modification of ear canal input impedance due to a mechanism of cochlear origin. The ratio of relative change in ear canal input impedance for a given relative change in cochlear input impedance will be a function of the degree of coupling achieved by the middle ear.

The observation of nonlinear ear canal impedance was made by interpreting the sound pressure response of the ear to a transient stimulus of varying level (Kemp, 1978). By stimulating the ear with a click it was shown that there existed a delayed, of the order of 6 ms, signal, typically of a duration of 15 ms, whose amplitude decreased proportionally with increasing stimulus level. Representing amplitude data as an input-output, growth function, the resulting function exhibited a compressive saturating nonlinearity.

It has been stated that the consequence of such a finding is that the ear canal impedance is nonlinear at low levels (Allen et al., 1995). However, this interpretation lacks accuracy. Under harmonic stimulation the level independent response (LIR) of the ear can be thought of as being augmented by a level dependent response (LDR) which is a function of the reemission mechanism. During stimulation of ~10 dB SL the gross response of the ear can be considered linear, since homogeneity is observed, indicating that both the LDR and LIR are operating in a linear regime (Shera and Zweig, 1993). At stimulus levels above 10 dB SL the LIR remains linear, whereas the reemission mechanism displays a saturating characteristic, resulting in a combined nonlinear gross ear response. With increasing stimulus level the LDR amplitude becomes proportional less, such that at stimulus levels above 60 dB SL the modification by the reemission mechanism becomes negligible. For stimulus levels above 60 dB SL the response of the ear is governed by the linear gross characteristic, up to the stimulus level which activates the acoustic reflex, approximately 80 dB SL. To summarise, two linear regimes can be identified in the gross acoustic response of the ear; one high regime ranging from approximately 60 dB SL to 80 dB SL, and another lower regime whose upper limit is roughly 20 dB SL (Shera and Zweig, 1993).
2.11 Method to measure ear canal input impedance

Occluding the ear canal by inserting an electroacoustic probe, consisting of a sound source and microphone, leads to a closed acoustic system. Under plane, one dimensional wave propagation, valid for frequencies less than 10 kHz, the system consists of a sound source coupled to an acoustic load. Under such a condition, the continuity of volume velocity holds true. The sound source is characterised in terms of its Thevenin equivalent, where $p_s$ and $Z_s$ are the perfect, zero impedance open circuit sound pressure source and series impedance respectively, collectively termed the Thevenin parameters, which are typically complex functions of frequency. Characterisation of a sound source in the frequency domain is usually facilitated by reference to the complex frequency response, and in the time domain by the impulse response. Assuming linearity, the Fourier transform pair provides the link between the impulse response and frequency response. However, both these measures not only carry information of the source but that of the load which is being driven; the impulse and frequency responses are load dependent. Calibration of the sound source as an equivalent Thevenin source, or its dual, referred to as a Norton source, enables the separation of the source from the load. The corresponding equivalent Norton source consists of a perfect, infinitely high impedance volume velocity source, $U_s = p_s / Z_s$, in parallel with an impedance, equal to $Z_s$. The acoustic load is expressed as a lumped impedance element, $Z_l$, equivalent to the acoustic input impedance of the load.
The relationship between the sound pressure measured in the probe-tip load plane, \( p_i \), to \( Z_i \), is given by the potential divider equation:

\[
p_i = p_s \frac{Z_i}{Z_i + Z_s}
\]  

(2.1)

Each variable is the complex function in frequency of the Fourier transform of the corresponding function in time. For the purposes of clarity, reference to the frequency dependence, i.e., \( e^{j\omega t} \), is dropped. Calibrating the sound source in terms of the Thevenin parameters and with \( p_i \) known allows for the estimation of the input impedance of acoustic loads, such as the ear.

2.12 Method to estimate Thevenin parameters

Lynch \textit{et al} (1994) estimated the Norton parameters of a sound source, intended for measurements in cat, by measuring the sound pressure at the entrance of a set of three reference loads; a small cavity, \( Z_{SC} \), a larger cavity, \( Z_{LC} \) and a long tube, \( Z_{LT} \), terminated by the tube's characteristic impedance, approximating to an infinitely long tube whose impedance is purely resistive. Each load had a differing input impedance characteristic with frequency, each optimal for a particular frequency range; \( Z_{SC} \) and \( Z_{LC} \) were used for frequencies below 0.1 kHz and \( Z_{SC} \) and \( Z_{LT} \) used above 3 kHz. An average of the two pairs were used for frequencies between 0.1 and 3 kHz. Such a
calibration approach has been used elsewhere Ravicz et al (1992). In order to assess the errors in the estimated Norton parameters the input impedance measures in two loads, different from the calibration set, were compared to model predictions. Accuracies of 10% in impedance magnitude and 0.02 periods in impedance argument for frequencies less than 12 kHz were quoted.

Allen (1985) describes a method to estimate the Thevenin parameters of a sound source and pressure microphone which make up a single probe assembly for the measurement of eardrum impedance in the cat. The probe assembly to be characterised contained an electret push-pull transducer connected, via a matching acoustic resistor in order to reduce reflections, to a uniform tube, 3.5 mm in diameter. Since the aim of acoustic calibration of the probe assembly is the estimation of two parameters, at least two independent responses are necessary. The complete sound source was coupled to a set of, four in this case, reference acoustic cavities, each with differing, but known, input impedances. Obtaining the pressure response between 0.2 and 33 kHz, in more than two loads over specifies the network, permitting the formalism of an over determined set of system equations which, with the additional degrees of freedom, allows for a more robust least mean squares approximation. To improve the accuracy of calibration, the mean of the residual error of the over determined equations across frequency, which reflects the mismatch between the sound pressure measurements and tube model data, was minimised by varying the length parameter in the tube model set.

Keefe et al (1992) presented a modification of Allen’s procedure described above for the frequency domain calibration of acoustic probe assemblies for impedance and energy reflectance measurements in human ear canals, across a frequency range between 0 to 10.7 kHz. In common with Allen’s method, measurements of pressure were matched to tube model data through the use of the least mean squares approximation. Six tubes of varying length and area were chosen for the calibration procedure. Inaccuracies in the estimates of lengths of tubes used during the sound pressure measurements lead to errors in the estimated Thevenin parameters. Keefe et al (1992) defined an error function which can be thought as the average normalised residual error of the over determined equations across the optimised frequency bandwidth. Such an error function differs from that used in Allen (1985) by an
additional weighting coefficient, which leads to an error function which is equally sensitive to maxima and minima in the impedance data.

An alternative approach to probe assembly calibration in the time domain is given in Keefe (1996) which uses a single long tube, of known dimensions, in order to separate forward travelling and backward travelling signals. Rather than express the response of the probe assembly in terms of Thevenin parameters, the calibration procedure estimates the reflection function due to the probe assembly source impedance.

2.13 Ear canal acoustic measurements

Keefe et al (1993) measured the acoustic input impedance of the human ear canal in subjects of various ages, at the tip of the probe assembly by using a Thevenin calibrated probe. A wide band chirp signal was generated at moderate stimulus levels. For all ages, from 1 month to adult, impedance magnitude curves were smooth and slowly varying.

Because of the frequency transforming property of the ear canal, since it essentially acts as a hard walled duct, the measure of input impedance is dependent upon probe location relative to the TM. A measure which can be derived from the input impedance, by making the assumption that the area of the canal is uniform along its length such that no significant energy is scattered back, is pressure reflectance. By taking the square of the magnitude of pressure reflectance one calculates energy reflectance, which for lossless transmission lines is independent of measurement position. Subtracting energy reflectance from one, gives the transmittance, a fundamental physical parameter in audition, since it is the ratio of transmitted energy to incident energy.

The behaviour of impedance, \( Z(\omega) \), and pressure reflectance, \( R(\omega) \), are dissimilar. For a simple system such as a lossless uniform duct, \( Z(\omega) \) changes rapidly with frequency and is singular\(^{13} \) at frequencies which correspond to standing waves, whereas

\(^{13}\) A function of a complex variable is singular when it ceases to be analytic. An analytic function is single valued and differentiable at all points in a region (O'Neil, 1987). For example, \( f(z) = z^2 \) is
$R(\omega)$ is analytic, defined as a vector rotating clockwise at a constant angular velocity against frequency.

Allen et al (1995) measured $R(\omega)$ as a function of stimulus intensity. The measured pressure reflectance exhibited a microstructure characteristic, with fluctuations more pronounced at lower stimulus intensities.

2.14 Summary

This chapter has described the nonlinear phenomenon of microstructure observable in acoustic and psychoacoustic auditory measurements and reviewed the studies that have addressed the underlying mechanism. Reference was made to physical concepts such as linearity, transmission line wave propagation and power gain, with respect to discussions on ear canal acoustics, cochlear modelling and the cochlear amplifier, respectively.

The data reviewed evinces that the nonlinear mechanism is of cochlear origin. The processing of high intensity signals accomplished primarily by the physiologically healthy cochlea is essentially different to that at stimulus intensities about absolute threshold; at high intensities energy propagates along the cochlea from base to apex, whereas low level stimuli initiate bi-directional energy flow, leading to the phenomenon of standing waves. A manifestation of cochlear standing waves is a spectral oscillation observable in acoustic and psychoacoustic parameters elicited at near threshold levels, generically termed auditory microstructure. Characteristics of the nonlinear cochlear mechanism that is integral to normal cochlear function can then be probed non-invasively with acoustic and psychoacoustic paradigms.
2.15 Scope of thesis

The primary purpose of this thesis is to report on an investigation to characterise auditory microstructure using acoustic parameters not commonly reported on in the literature. The motivation of using acoustic parameters, such as pressure reflectance and power, is to provide a more relevant description of the phenomenon of auditory microstructure than the conventional measure of sound pressure. By modelling the peripheral auditory system using reflectance and power descriptions, the nature of the cochlear travelling wave interaction may be determined.

Knowledge of the sound pressure across a known acoustic impedance element enables the determination of the acoustic power incident upon the element. Since the magnitude squared of \( R(\omega) \) is equal to the power reflectance, i.e., the ratio of reflected power to incident power, separation of the incident power into reflected and absorbed constituents is possible. Power is a fundamental descriptor for any receptor system and unlike sound pressure, which is a vector quantity, power is scalar and as such is not prone to phasic interactions which lead to the phenomenon of interference.

This thesis will report an ameliorated Thevenin probe assembly calibration procedure which allows for impedance measurements derived from CTASP in adult human ear canals. By using a vector subtraction method, analysis of the additional sound pressure component at low levels, with reference to impedance and reflectance parameters, will be made. Analysis on the desideratum of separating acoustic power measured in the ear canal into incident, reflected and transmitted constituents will be achieved by use of scattering parameters. Absolute power absorbed at the tympanic membrane at threshold is shown.
2.16 Outline of thesis

In chapter 3 the theory of the Thevenin probe assembly calibration method will be developed. This method first described in Allen (1985), is a frequency domain method that employs a calibration tube set. Here, the theory will be extended to include measures of the confidence in the estimated parameters.

Chapter 4 describes the method used for Thevenin probe calibration. The choice of the calibration tube set, mathematical modelling of acoustic loads and the effects of evanescent modes upon sound pressure measurements will be made.

The Thevenin parameters of the probe are given in chapter 5. An investigation of the effect of port tube length and inclusion of acoustic resistor in the probe construction upon the Thevenin parameters is detailed. Measurements of the acoustic input impedance, using the calibrated probe, of a hard walled rigidly terminated tube, a Brüel & Kjær ear simulator and adult ear canals are shown.

Raw continuous tonal ear canal sound pressure measurements are detailed in chapter 6 together with results of the variation in the response with level. Discussion will focus on the influence of the electroacoustic probe used to deliver and sample the sound within the closed ear canal. The nature of, and errors in, the subtracted sound pressure, identified as the stimulus frequency otoacoustic emission, will be discussed.

In chapter 7 three acoustic parameters are derived from the raw ear canal sound pressure; ear canal input impedance, pressure reflectance and power flow. The three constituent parts of ear canal power flow, namely power incident, reflected and absorbed at the measurement point, will be presented. It is shown that all three measures exhibit a microstructure characteristic.

Chapter 8 presents cochlear input impedance and pressure reflectance inferred from the empirical data shown in chapters 6 and 7 through the use of an analogue network model of the peripheral auditory system. It is argued that the origin of microstructure can arise through the phasic interaction of forward and backward travelling waves. A simple
form of cochlear pressure reflectance against frequency, which carries information on the phasic interaction, explains the microstructure observed in ear canal parameters.

Chapter 9 describes the methodology and presents the results of auditory thresholds determined at small frequency intervals. The issue of modification in coupling between the sound source and ear and its role in threshold microstructure is address with the formulation of a phenomenological model of power flow within the peripheral auditory system. A non-uniform transmission line model of the cochlea will be implemented to illustrate the effect of bi-directional energy flow on the distribution of energy along the cochlea.

Conclusions are made in chapter 10 together with suggestions for further research.

In chapter 11 all references made in the thesis are listed in alphabetical order of the first author, together with title, journal and year of publication.
3.1 Formalism of the least mean squares solution of an overdetermined set of system equations

This chapter will detail the theory of the Thevenin calibration of electroacoustic probe assemblies used for the delivery and measurement of sound pressures in the human ear canal. Although the theory makes use of the least mean squares solution of an overdetermined set of system equations which can be found in linear algebra textbooks (Fraleigh and Beauregard, 1995) and applied texts (Gujarati, 1988), coverage is limited to real variables. In the application of acoustic probe calibration the variables are invariably complex and it is for this reason that the least mean squares solution is developed below.

3.2 Definition of an over-determined system

The \(i\)th generalised system equation of an over-determined system of \(k\) linear equations of two parameters, is given by:-

\[
x_1 a_{ui} + x_2 a_{2i} = b_i
\]

where:

\[
k > 2;
\]
\(a_{ui}\) and \(a_{2i}\) are explanatory variables;
\(b_i\) is the dependant variable;
\(x_1\) and \(x_2\) are the model parameters.

All variables and parameters are considered to be complex to keep the formalism as general as possible.

Rearranging the potential divider equation (equation 2.1) gives:-
Chapter 3: Probe assembly calibration: Theory

\[ Z_i p_s - p_s Z_s = Z_i p_i \]  \hspace{1cm} (3.2)

i.e.,

\[ a_{ii} = Z_i, \text{ for } i=1..k; \]
\[ a_{2i} = -p_i, \text{ for } i=1..k; \]
\[ b_i = Z_i p_i, \text{ for } i=1..k; \]

We then have the system equations as defined in Allen (1985) and Keefe et al (1992), where \( x_1 = p_s \) and \( x_2 = Z_s \). Allen (1985) used four tubes, i.e., \( k = 4 \), Keefe et al (1992) \( k = 6 \). For the purposes of clarity, the frequency dependence on each pressure variable, \( e^{j\omega t} \), has not been made explicit. To clarify, the system equations are valid at one particular frequency only, and that estimation of the model parameters at another frequency requires the appropriate set of system equations.

3.3 Matrix representation of overdetermined system equations

Therefore, in matrix form, an over-determined system of \( k \) linear equations, where \( k > 2 \), can be written as:

\[
\begin{bmatrix}
  b_1 \\
  b_2 \\
  \vdots \\
  b_k
\end{bmatrix} =
\begin{bmatrix}
  a_{11} & a_{21} \\
  a_{12} & a_{22} \\
  \vdots & \vdots \\
  a_{1k} & a_{2k}
\end{bmatrix}
\begin{bmatrix}
  x_1 \\
  x_2
\end{bmatrix}
\]

\[
b = A x
\]  \hspace{1cm} (3.3)

where:-

\( b \) is a column vector of length \( k \) of dependent variables \( b_i \);

\( A \) is a \( k \times 2 \) matrix;

\( x \) is a column vector of length 2 of the unknown parameters \( x_1 \) and \( x_2 \);
With the error defined as the difference between the model estimate and measurement data, i.e., \( e_i = b_i - (x_1a_{1i} + x_2a_{2i}) \) then it follows \( b = Ax + e \), where \( e \) is a column vector of length \( k \), of error terms, \( e_i \).

### 3.4 Principle of least mean squares

The least mean squares (LMS) solution is that solution which minimises the sum of the magnitude square of the error terms, which can be thought of as minimising the energy in the error vector. For complex variables, the function to minimise is given by:

\[
J_{\text{com}} = \sum_{i=1}^{k} |e_i|^2
\]

In matrix form \( J_{\text{com}} \) is found by noting that the Euclidean inner product of \( e \), \( \langle e, e \rangle \), is given by:

\[
\langle e, e \rangle = e^T \cdot e
\]

where:

\( e^T \) is the operation of complex conjugate transposition of \( e \).

The LMS solution of an overdetermined system can be arrived at by an approach aided by a graphical construction (Fraleigh and Beauregard, 1995). Another solution employs matrix differentiation (Gujarati, 1988). Both Fraleigh and Beauregard (1995) and Gujarati (1988) provide the solution for real variables, whereas below the solution is developed for complex variables.

### 3.5 Graphical construction approach to method of least mean squares

Define \( C^k \) as the set of all complex numbers in \( k \)-dimensional space. We define \( W = \text{sp}(a_1, a_2) \) to be a subspace of \( C^2 \), i.e., \( W \) is a complex plane. Set \( x = x_{\text{min}} \) as the optimal solution vector.
We use the principle of orthogonal projection in order to decompose a vector, \( \mathbf{b} \), into two orthogonal vectors, such that \( \mathbf{b} = \mathbf{b}_n + \mathbf{b}_n\perp \). The vector which minimises the length \( \| \mathbf{b} - \mathbf{A} \mathbf{x}_\text{min} \| \) is the projection of \( \mathbf{b} \) on \( W \); i.e., \( \mathbf{b}_n = \mathbf{A} \mathbf{x}_\text{min} \). The vector \( \mathbf{b} - \mathbf{A} \mathbf{x} \) is perpendicular to every vector in \( W \), in other words \( \mathbf{b}_n \perp = \mathbf{b} - \mathbf{A} \mathbf{x} \); the dot product of the vectors \( \mathbf{A} \mathbf{x} \) and \( \mathbf{b} - \mathbf{A} \mathbf{x} \) is zero for all vectors \( \mathbf{x} \).

\[
(\mathbf{A} \mathbf{x}_\text{min})^\top (\mathbf{b} - \mathbf{A} \mathbf{x}_\text{min}) = 0 \tag{3.6}
\]

Rearranging the l.h.s. gives:-
Chapter 3: Probe assembly calibration: Theory

\[
\mathbf{x}_{\text{min}}^\top (\mathbf{A}^\top \mathbf{b} - \mathbf{A}^\top \mathbf{Ax}_{\text{min}}) = 0 \tag{3.7}
\]

For \( \mathbf{x}_{\text{min}} \neq 0 \), the nontrivial case, it follows that:

\[
\mathbf{x}_{\text{min}} = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{b} \tag{3.8}
\]

The general LMS solution is then given by:

\[
\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{1}{\Delta} \begin{bmatrix} \sum a_{2i}^2 - \sum a_{ii} a_{2i} \\ \sum a_{ii}^2 \end{bmatrix} \tag{3.9a}
\]

where:

\[
\Delta = \sum |a_{ii}|^2 \sum |a_{2i}|^2 - \sum a_{ii} a_{2i} \sum a_{2i} a_{ii} \tag{3.9b}
\]

The preceding solution is employed in many studies that aim to calibrate electroacoustic probes (Allen, 1985; Keefe et al, 1992). However, further information relating to the uncertainty of each of the two Thevenin parameters can be obtained through the determination of the variance-covariance matrix. Knowledge of the variances allow for the process of statistical inference.

### 3.6 Variance and Standard errors

The variance in the data, \( \sigma^2 \), is given by:

\[
\sigma^2 = \frac{\mathbf{e}^\top \mathbf{e}}{k - 2} \tag{3.10}
\]

It is noted that for six independent system equations and two model parameters leads to four degrees of freedom. For the case of two independent system equations, as used in
the methods employed by Lynch et al (1994) and Ravicz et al (1992), determining the variance in the data is not possible, since two degrees of freedom are required in the estimation of $x_j$ and $x_k$. In the two studies mentioned above, the errors on the Thevenin parameters were estimated by comparing the experimental and model input impedance of known acoustic loads that did not make up the calibration set.

The two-by-two variance-covariance matrix, \textbf{var-cov}, is then given by (Gujarati, 1988):-

$$
\text{var-cov} = \frac{\sigma^2}{\Delta} \begin{bmatrix}
\sum |a_{2j}|^2 & -\sum a_{2j}a_{2l} \\
-\sum a_{2l}a_{2j} & \sum |a_{2l}|^2
\end{bmatrix} = \begin{bmatrix}
\sigma^2_{x_j} & \text{cov}(\sigma^2_{x_j}, \sigma^2_{x_k}) \\
\text{cov}(\sigma^2_{x_j}, \sigma^2_{x_k}) & \sigma^2_{x_k}
\end{bmatrix}
$$

(3.11)

The positive diagonal elements of \textbf{var-cov} give the variances of $x_j$ and $x_k$. It follows that the positive square roots of $\sigma^2_{x_j}$ and $\sigma^2_{x_k}$ give the corresponding standard deviations, with the estimated standard error of the mean given by:-

$$
\text{sem}_{x_j} = \frac{\sqrt{\sigma^2_{x_j}}}{\sqrt{k-2}}
$$

(3.12)

The 95% confidence interval, CI95%, with t-score equal to 2.78 for $df=4$, is then given by:-

$$
\text{CI95%}_{x_j} = x_j \pm 2.78 \text{sem}_{x_j}
$$

(3.13)

Covariances are given by the negative diagonal elements, $\text{cov}(\sigma^2_{x_j}, \sigma^2_{x_k})$ and $\text{cov}(\sigma^2_{x_k}, \sigma^2_{x_j})$. 
3.7 Minimisation of error function

Errors in the measurement of the tube length and radius leads to a difference between the model predictions and measured parameters. Amelioration of the Thevenin parameter calibration procedure is thus achieved by minimising the mismatch, i.e., an error function, by finding the optimal length and radii parameter values. The principal feature of such an error function is that it is a single real valued measure of the mismatch between the model predictions and measured parameters, not only across tubes at one particular frequency, but across the frequency bandwidth of interest. Defining an error function is only possible for the overdetermined case, since two degrees of freedom are required to describe $p_x$ and $Z_r$. For example, the calibration routine delineated in Lynch et al (1994) does not allow for the use of an error function and therefore improvement in the estimation, in this case, of the Norton parameters cannot be achieved.

Allen (1985) defined an error function, $E_T$, equal to the sum of $f_{com}$ over the frequency range of interest, $\omega$, to $\omega_f$, where $\omega_i$ is the $i^{th}$ frequency:

$$E_T = \sum_{\omega = \omega_1}^{\omega_2} f_{com}(\omega)$$

(3.14)

Keefe et al (1992) defined an average normalised error function, $N_T$, in which $f_{com}$ is divided by the magnitude square of the vector $b$:

$$N_T = \frac{1}{\omega_2 - \omega_1 + 1} \sum_{\omega = \omega_1}^{\omega_2} \left( \frac{f_{com}(\omega)}{\sum_{i=1}^{4} |b_i(\omega)|^2} \right)$$

(3.15)

Common to both $E_T$ and $N_T$ is the fact that the error functions describe a product of pressure and impedance. Minimising such a quantity, whose dimensions are not intuitive, appears not to be physically meaningful.
We define a new error function, $L_T$, in which the mean of the average of the normalised errors in $p_S$ and $Z_S$ across the bandwidth is calculated:

$$L_T = \frac{1}{\omega_2 - \omega_1} \sum_{\omega = \omega_1}^{\omega_2} \left( \left| \text{se}_{p_S}(\omega)/p_S(\omega) \right| + \left| \text{se}_{Z_S}(\omega)/Z_S(\omega) \right| \right)/2 \quad (3.16)$$

In common with Keefe et al (1992) we use Powell’s $k$-dimensional minimisation routine which supplies Brent’s one dimensional minimisation procedure with linear directions which have good convergence properties. All routines are given as a software implementation in Press et al (1994).
Chapter 4: Probe assembly calibration: Methods

4.1 Probe construction

The probe assembly to be calibrated was an Otodynamics BS probe, consisting of a Knowles BP1712 miniature loudspeaker and Knowles EM3046 miniature microphone. Tubing of length 19 mm and 2 mm ID is coupled to the port of the loudspeaker via a 2.2 K MKS acoustic ohm resistor. The microphone is ported by tubing of length 8 mm and 1 mm ID. Both the loudspeaker and microphone, together with a narrow diameter static pressure equalisation tube, are encased in a hard capsule. The significant difference between the BS probe and the probe assemblies used in previous studies (Keefe et al, 1993; Voss and Allen, 1994) is that in the former, the microphone port does not protrude from the solid probe tip, leading to the BS probe being an example of a flush ported probe construction. Consequently, with reference to following sections, such a flush ported probe can lead to the sampling of the evanescent\(^1\) wave in certain acoustic loads.

4.2 Choice of reference loads

To a first approximation the sound source within the probe, when situated in cavities similar in dimensions to that of an occluded adult ear canal, acts as a constant volume velocity source as opposed to a constant pressure source. Under such a drive condition, sound pressure magnitude developed across a load is proportional to the input impedance magnitude of that load. Suitable acoustic loads should have dimensions that are relevant to the measurement in hand, and a geometry that is easily constructed. Additionally, the acoustic loads should give rise to measurable pressures, i.e., sound pressure levels that span the frequency range of interest and which are above the system noise, as well as defined acoustic impedances which can be simply modelled mathematically.

\(^{1}\) As the term suggests, the evanescent wave is ephemeral, but rather than being transitory in the temporal domain, the pressure amplitude is quickly fading in a spatial sense.
The use of acoustic resistors and small short holes to provide resistance and mass in the acoustic load was avoided for two reasons. Firstly, to determine mathematically the acoustics requires factors, referred to as Karal elements, to correct for the effect of the rapid change in area between the tube and dampers, which can be thought of an additional series mass. The determination of the Karal correction factors appears to be rather heuristic (Voss and Allen, 1994). Secondly, small holes and acoustic dampers made up of fine mesh, over time become fouled with dust and debris which can greatly affect the acoustical characteristics.

Taking into account the above criteria, the convenient choice of cavities is a set of cylindrical, hard walled, rigidly terminated tubes. The tube cavities then differed in length and diameter.

The choice of length of each tube is consequential because large errors in the estimated Thevenin parameters occur when the input impedance of the set of loads are comparable. If this is the case, the points of the system equations all lie in close proximity on the complex plane, resulting in a LMS solution which is very sensitive to small errors on each point. The aim was to choose a set of lengths which lead to tube input impedances which differed maximally across the measurement bandwidth.

*Table 4.1: Dimensions of tubes used in calibration routine.*

<table>
<thead>
<tr>
<th>Tube number</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length (cm)</td>
<td>8</td>
<td>4.5</td>
<td>4</td>
<td>3.5</td>
<td>3</td>
<td>2.4</td>
</tr>
<tr>
<td>Radius (cm)</td>
<td>0.425</td>
<td>0.375</td>
<td>0.425</td>
<td>0.45</td>
<td>0.35</td>
<td>0.325</td>
</tr>
</tbody>
</table>
4.3 Mathematical modelling of acoustic loads

The following treatment is valid for one dimensional plane wave propagation, a qualification which has important consequences in the subsequent discussion in section 4.6.

Each load is modelled by a transmission matrix approach. The transmission (ABCD) matrix for a circular tube, length \( l \), of series impedance per unit length, \( Z \), and shunt admittance per unit length, \( Y \), is given by (Lampton, 1978; Keefe, 1984)

\[
\begin{pmatrix}
    p_i \\
    U_i
\end{pmatrix} =
\begin{pmatrix}
    \cosh(\Gamma l) & Z_o \sinh(\Gamma l) \\
    1/Z_o \sinh(\Gamma l) & \cosh(\Gamma l)
\end{pmatrix}
\begin{pmatrix}
    p_i \\
    U_i
\end{pmatrix}
\]  
(4.1)

where \( Z_o \) is the lossy\(^2 \) characteristic impedance:

\[
Z_o = \sqrt{\frac{Z}{Y}}
\]  
(4.2)

and \( \Gamma \) is the propagation wavenumber:

\[
\Gamma = \sqrt{ZY}
\]  
(4.3)

\( p \) and \( U \) correspond to sound pressure and volume velocity respectively, with the subscripts, \( i \) and \( t \), referring to two different points along the tube, the input and terminator respectively. Defining the terminating impedance as \( Z_t = p_t/U_t \), gives the input impedance, \( Z_i \):

---

\(^2\) Lossy implies energy dissipation along the length of transmission line, i.e., non-zero attenuation per unit length, in contrast to a lossless transmission line which exhibits no energy dissipation between input and terminator. It is found that the attenuation per unit length for tubes of dimensions of the order of the human adult ear canal is small, roughly 0.06 dB/cm. However, during the Thevenin calibration routine it is necessary to model the acoustic characteristics of the calibration tubes as defined above, and therefore, the term lossy is used to maintain accuracy.
For a rigidly terminated cylindrical tube, where $Z_i \to \infty$, leads to:

$$Z_i = Z_o \coth(\Gamma l) \quad (4.5)$$

The preceding treatment is valid for both isothermal and non-isothermal boundary conditions, i.e. temperature changes of the inner tube wall during the oscillatory cycle. For the case when the inner tube wall can be considered isothermal, which will be adequate for our application, the solid-gas interface has zero capacity for the transfer of heat. Kirchhoff’s (1868) solution with isothermal boundary conditions for the acoustic propagation in a rigid cylindrical tube includes thermal and viscous effects of the fluid. Keefe (1984) provides approximations of the Kirchoff solution for the determination of $Z_o$, $\Gamma$, $Z$ and $Y$ and notes that the fluid thermal losses are carried in $Y$, while $Z$ is a function of the viscous losses. The combined viscothermal losses constitute the real part of $Z_i$ and determine the dissipation of energy in the acoustic load.

Table 4.2: Thermodynamic constants used in the modelling of acoustic loads. Accurate to $\pm10\%$ at temperature of 300 °K (26.85 °C). From Keefe (1984).

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Equilibrium gas density $\rho$</td>
<td>1.1769 kg m$^{-3}$</td>
</tr>
<tr>
<td>Speed of sound in free space $c$</td>
<td>347.23 m s$^{-1}$</td>
</tr>
</tbody>
</table>
Figure 4.1: Input impedance of each tube in the calibration tube set. The parameter in each panel refers to the tube listed in Table 4.1.
From figure 4.1 the magnitude of the input impedance of a hard walled rigidly terminated tube, when plotted logarithmically, is symmetric about a mean value equal to its $|Z_d|$. The ratio of $|Z_{\text{max}}|$ to $|Z_{\text{min}}|$ is of the order of 70 dB.

### 4.4 Construction of reference loads

Transparent cast acrylic rod of o/d. 20 mm was utilised. Tubes were turned up with each end plate attached by tape to the barrel. To ensure a good hermetic seal with the two end plates, petroleum jelly was applied at each end of the tube and secured with insulation tape. Table 4.1 lists the dimensions of the six tubes. The model input impedance of each tube is shown in figure 4.1

### 4.5 Stimulus generation and data acquisition

All measurements during the Thevenin probe assembly calibration routine were implemented on a Brüel and Kjær Type 2035 dual channel signal analyser. The analyser performed three tasks. Firstly, to generate the stimulus. Secondly, to calculate the Fourier transform between the measured complex pressure and voltage drive input. Thirdly, to save magnitude and phase responses to a 3.5" floppy disk.

The stimulus generated was a periodic multisine (chirp) signal, which in the frequency domain has spectral lines of equal magnitude, and a phase characteristic which increases with the square of frequency. A multisine signal was chosen due to a relatively low crest factor, i.e., the ratio of peak to rms level, ensuring that nonlinearity from the loudspeaker was minimised. The crest factor of a multisine signal is 1.65. Additionally, no leakage occurs when used in conjunction with rectangular windowing, reducing artefacts during data analysis. Synchronous time averaging which efficiently reduces random noise can be employed due to the periodic nature of the signal. The noise across the bandwidth during a typical measurement was less than 0 dB SPL after 256 averages over a time course of approximately 35 s (see figure 5.2).
4.6 Propagating and evanescent modes

When a sound source generates an acoustic field the resulting pressure fluctuations can be thought of as a sum of an infinite number of modes (Morse and Ingard, 1968). This observation reflects that, generally, there is more than one, non-zero, solution of the wave equation which satisfies a set of boundary conditions. Each solution is termed an eigenfunction\(^3\) and has an associated eigenfrequency.

Consider the case of harmonic stimulation of air within a hard walled circular tube of radius \(a\), driven by a piston at the \(z = 0\) boundary. Referring to the usual cylindrical co-ordinates, \(r, \theta\) and \(z\), the radial, azimuth and axial co-ordinates respectively, the eigenfunction, i.e., the solution for sound pressure, is given by (Kinsler et al, 1982):

\[
P_{mn} = A_{mn} J_m[k_{mn} r] \cos[m \theta] \exp[j(\omega t - k_z z)]
\]  

(4.6)

where \(m\) and \(n\) are natural numbers and refer to the mode number; \(J_m\), is the \(m^{th}\) order Bessel function of the first kind; \(j = (-1)^{1/2}\); \(\omega = 2\pi f\) where \(f\) is the driving frequency; \(k_z\) is the axial wavenumber, which determines the nature of propagation, and is given by:-

\[
k_z^2 = k^2 - k_{mn}^2
\]  

(4.7)

where the propagating wavenumber, \(k = \omega / c\), and \(c\) is the speed of sound. \(k_{mn}\) is the modal wavenumber given by:-

\[
k_{mn} = \mu_{mn} / a
\]  

(4.8)

where \(\mu_{mn}\) is the \((n+1)^{th}\) zero of \(dJ_m[x] / dx\), i.e., inflection point of \(J_m[x]\).

\(^3\) Prefix from the German eigen, signifying own, characteristic.
Table 4.3: Table of inflection points of the m\textsuperscript{th} order Bessel function of the first kind.

<table>
<thead>
<tr>
<th>m\textsuperscript{th} order</th>
<th>i\textsuperscript{th} inflection point</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0 3.83 7.02</td>
</tr>
<tr>
<td>1</td>
<td>1.84 5.33 8.54</td>
</tr>
<tr>
<td>2</td>
<td>3.05 6.71 9.97</td>
</tr>
</tbody>
</table>

For the (0,0) mode, $\mu_{00} = 0$, therefore $k_{00} = 0$, leading to $k_z = \omega / c$. Since $k_z$ is real, the (0,0) mode propagates with speed equal to $c$, for $\omega > 0$. The (0,0) mode is therefore an example of a propagating mode. In addition, because the (0,0) mode is completely described by one spatial variable, the axial co-ordinate, $z$, this propagating mode is an example of a plane wave. Here, the phase of neighbouring wavefronts relative to the stimulus varies with distance, but is related through $k$ in a linear manner.

However, for the (1,0) mode, $k_{10} \approx 1.84$, such that $k_z$ is imaginary for low frequencies but real for frequencies above a point termed the modal cut-off frequency $f_{\text{mn}}$. The cut-off frequency for the (1,0) mode, which is the lowest $f_{\text{mn}}$, is given by $f_{1,0} \approx 101/a$ (Kinsler et al, 1982). For $a=4.5$ mm the lowest cut-off frequency is then 22.4 kHz.

When $f > f_{\text{mn}}$, a propagating mode exists, resulting in a non-planar wave due to the dependence upon more than one spatial co-ordinate.

When the wavenumber $k_z$ is purely imaginary, that is when $f < f_{\text{mn}}$ then the mode is termed (non-propagating) evanescent\(^4\). The consequence of an evanescent mode is an acoustic wave in the medium which has two distinct properties from the plane wave case. Firstly, points in close proximity to the source vibrate in synchrony, in other words the displacement of the medium is in phase with the displacement of the sound source. Under such a condition, the wave velocity is in quadrature to the displacement.

\(^4\) Such modes turn up in optics and quantum mechanics and give rise to the phenomenon of tunnelling.
Chapter 4: Probe assembly calibration: Methods

This leads to the observation that the disturbance has zero capacity for energy dissipation. Rather than being dissipated, the energy is stored, as if the medium acted as a purely reactive impedance. Secondly, a snapshot in time of the amplitude against z is not sinusoidal, as for plane waves, but rather decays exponentially.

Table 4.4 lists the, customary, 1/e distance, $\lambda_{mn}$, which is the axial length in which the evanescent mode amplitude falls by a factor of approximately 0.37. Rearranging equation 4.1 gives $k_z = (\mu_{mn} / a)(1 - (ka / \mu_{mn})^2)^{1/2}$. When $ka << \mu_{mn}$, termed the low frequency approximation, the wavenumber approximates to $k_z \approx (\mu_{mn} / a)$ and is independent of frequency. Therefore, $\lambda_{mn}$ is approximately equal to $(a / \mu_{mn})$ (Keefe and Benade, 1981). For the worst case, that is in the (1,0) mode, $\mu_{10} = 1.84$, and for $f=10$ kHz and $a=4.5$ mm, the largest radius in the calibration tube set, the error in $\lambda_{mn}$ for the low frequency approximation is -11%.

Table 4.4: 1/e distances for circular tube, $a=4.5$ mm. Low frequency approximation used. Distance in mm.

<table>
<thead>
<tr>
<th>$m^0$ mode</th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>N/A</td>
<td>1.17</td>
<td>0.64</td>
</tr>
<tr>
<td>1</td>
<td>2.25</td>
<td>0.84</td>
<td>0.53</td>
</tr>
<tr>
<td>2</td>
<td>1.48</td>
<td>0.67</td>
<td>0.45</td>
</tr>
</tbody>
</table>

The discussion so far has been valid for the origin of excitation to be an oscillatory boundary, equivalent to a piston source. Consider the case where the source is taken to act at a point, $(r_s)$. A microphone situated at $(r_1, \theta_1, z_1)$ samples the pressure disturbance. The preceding treatment of propagating and evanescent modes equally applies to this case. Additionally, the following discussion is valid for one dimensional acoustic wave propagation only, in that all eigenfunctions are evanescent modes, except for the (0,0) mode which is a propagating mode, planar in dispersion.
The sum of the evanescent modes is given by a single real valued non-dimensional parameter, termed the evanescent mode factor, $\varepsilon$, which as Keefe and Benade (1981) note is a function of the configuration of the sound source relative to the measurement point.

We define an evanescent wave as that disturbance which is a result of all contributing non-propagating modes.

The presence of the evanescent pressure wave is modelled as a result of an additional impedance lumped element, $Z_{ev}$, in series with the input impedance of the acoustic load, $Z_{in}$, which is dependent upon the plane wave only, as illustrated in figure 4.2.

**Figure 4.2:** Network analogue illustrating the effect of the additional evanescent term as an additional impedance, $Z_{ev}$, which is purely reactive, assuming one dimensional wave propagation.

![Network analogue](image)

Under harmonic stimulation, $e^{j\omega t}$, the evanescent dependant impedance element, $Z_{ev}$, is then given by $jR_o kac[r_o, r_1, \theta_1, z_1]$ (Keefe and Benade, 1981; Brass and Locke, 1997), where $R_o$ is the lossless characteristic impedance given by $pc/\pi a^2$. $Z_{ev}$ is, by virtue of the $90^\circ$ operator, $j$, purely reactive and exhibits an argument which is governed by the sign of the evanescent mode factor; $+90^\circ$ for $\varepsilon > 0$, and $-90^\circ$, for $\varepsilon < 0$. As stated previously, the pressure disturbance due to the evanescent wave dissipates zero energy, compatible with the inclusion of a reactive impedance. Also, the frequency behaviour
of \( Z_{ev} \) is mass-like due to the proportional dependence upon \( \omega \), so that \( |Z_{ev}| \) increases at 6 dB/Oct.

Consider the situation in which in cylindrical co-ordinates, the loudspeaker port is situated at \((r_0 = R)\) and the microphone located at \((r_1 = R, \theta_1 = \pi, z_1 = 0)\), which roughly corresponds to the geometry of the BS probe used in this study. From Brass and Locke (1997) this configuration approximately corresponds to the \( z = 0.025a \) curve in panel D in their figure 5. When the radius of the acoustic load is \( a = 4.5 \) mm and \( R = 1.5 \) mm then \( \varepsilon \approx -0.3 \).

### 4.7 The effect of the evanescent wave upon acoustic measurements

The effect of the evanescent wave upon acoustic measurements has been investigated in couplers commonly used in audiology (Burkard and Sachs, 1977), expressed mathematically (Morse and Ingard, 1968; Keefe and Benade, 1981) and examined in human ear canals (Brass and Locke, 1997).

To summarise the effect of the evanescent wave, at frequencies where the magnitude of the input impedance of the load is small the effect is greatest. For hard walled rigidly terminated circular tubes, of radii comparable to the dimensions of the probe used for ear canal measurements, the acoustic input impedance magnitude minima rarely is less than 0.1 M MKS acoustic ohm\(^5\) (see figure 4.1).

---

\(^5\) The S.I. derived unit of acoustic impedance is MKS (metre-kilogram-second) acoustic ohm corresponding to base units of kg m\(^{-4}\) s\(^{-1}\). In the past, many studies quantified acoustic impedance in the non-S.I. units of CGS (centimetre-gram-second) acoustic ohm. To obtain MKS acoustic ohm multiply CGS acoustic ohm by 10\(^5\).
Chapter 5: Probe assembly calibration: Results

5.1 Sound pressure measurements

The experimental set-up used for the determination of the Thevenin parameters is illustrated in figure 5.1. Under electrical stimulation of the loudspeaker with a periodic (chirp) signal, at a level of -35 dBV and a bandwidth of 12.8 kHz, an acoustic wave at the probe tip is launched onto the acoustic load. As stated in section 4.6, the acoustic wave is made up of the plane, and evanescent, waves.

The amplitude of each evanescent mode diminishes exponentially with axial distance, such that the effect of the evanescent wave is negligible at the termination of the tube (see Table 4.4). The Brüel and Kjær probe tube microphone Type 4182 attached to the venting tube attachment Type DB2930 via 25 mm of flexible tubing Type AF0576, allows for the sound pressure to be sampled at a location flush with the interior side of the terminating disc.

*Figure 5.1: Schematic of the stimulus delivery and sound pressure measurement system used for Thevenin electroacoustic probe calibration.*
Positioning the microphone at the termination of the tube requires a twofold modification in the calibration procedure to that used in previous studies (Allen, 1985; Keefe et al, 1992).

Firstly, with respect to the model representation of the tube, the termination impedance cannot be taken as infinitely high, i.e., rigid, but rather equal to the input impedance of the probe tube microphone. Dimensionally, the radius of the probe tube microphone is small relative to the tube radius, such that the area ratio of the probe tube microphone to the tube is approximately 0.025. Additionally, the design of the probe tube microphone is such that in order to minimise reflections the microphone incorporates an impedance matching cavity (Brüel and Kjær, 1990). Noting these two points, the input impedance of the probe tube microphone was modelled as mainly resistive with a value roughly equal to the probe tube's characteristic impedance; leading to a figure of 800 M MKS acoustic ohm. By using a transmission matrix approach, the input impedance of the tube, calculated for a terminator equal to 800 M MKS acoustic ohm, is symbolically equivalent to $Z_i$ as given in Keefe et al (1992).

Secondly, to obtain the sound pressure at the probe assembly tip, the sampled sound pressure at the terminator is transformed by the two port network which represents the acoustics of the model tube. The transformed pressure is then symbolically equivalent to $p_i$ as expressed in Keefe et al (1992).

Sampling the sound pressure at the high impedance termination, rather than the entrance of the tube, ameliorates the signal to noise ratio (SNR) in the measurement. The magnitude of the input impedance at the open end of a rigidly terminated tube against frequency is characterised by a series of consecutive valleys and peaks. The ratio of impedance magnitude maxima to minima, $Z_{RAT}$, for a hard walled rigidly terminated tube is predominately a function of the tube radius. However, for radii comparable to that used in the calibration $Z_{RAT}$ is roughly 70 dB (see figure 5.1). At frequencies

---

1 The Brüel and Kjær manual (1990) states that the impedance presented by the probe tube microphone is very high, such that "in the frequency range above 50 Hz and for volumes greater than 1 cm$^3$ the influence of the loading is negligible".
which correspond to impedance magnitude minima the source is acting as a near perfect constant volume velocity source, manifesting a proportionality between impedance and sound pressure; the sound pressure magnitude exhibits minima at frequencies coincident with the impedance magnitude minima. It follows that the sound pressure sampled at frequencies coinciding with impedance magnitude minima is contaminated proportionally more by noise than at frequencies away from impedance magnitude minima, leading to a lower SNR at impedance troughs. Conversely, the sound pressure magnitude at the termination is relatively flat against frequency, since the impedance is dominated by the high impedance terminator, the consequence is that the SNR is approximately constant across the measurement bandwidth.

**Figure 5.2: Difference in sampling at the probe tip compared to the terminating disc.** Measurement in calibration tube number 1, i.e., a rigidly terminated hard walled tube of length 8 cm, radius 0.425 cm. Reference is made to the resulting SNR at both measurement sites. Thin line is the noise floor, i.e., the output of the microphone when no stimulus is presented.
From figure 5.2 the SNR at certain frequencies is as small as 10 dB when the measurement point is at the probe tip, but greater than 30 dB at the terminating disc.

### 5.2 Measuring microphone frequency response

Data throughout this report is given such that the influence of the measurement system is transparent, i.e., the transfer characteristic of the measuring microphone *in situ* is subtracted. Two microphones were predominately used in this study; a B&K probe tube microphone Type 4182, for use in the calibration procedure; and a miniature electret microphone in the Otodynamics BS probe for use in ear canal acoustic measurements. Determination of the frequency response of each microphone under a white noise free-field condition was achieved by comparing the microphone output to that of a B&K Type 4134 1/2” calibrated reference pressure microphone of a known transfer characteristic.
Figure 5.3: Frequency response of two microphones under white noise free-field conditions. The Brüel and Kjær probe tube microphone Type 4182 with 25 mm flexible tubing (solid line) and the microphone unit located in the Otodynamics BS probe (dashed line). The magnitude response, top panel, is expressed as \(20 \log_{10}(V_m/V_o)\), where \(V_o\) is equivalent to a sensitivity of 1 V/Pa. The bottom panel displays the phase response referenced to the incident sound pressure, in units of degrees. Data points every 16 Hz with straight line interpolation.
The frequency response of the Brüel and Kjær probe tube microphone Type 4182 (solid line) displays a smooth slowly varying characteristic which reflects the internal impedance matching element. When plotted on a logarithmic frequency scale the magnitude of the frequency response manifests a low pass characteristic; 6 dB/octave first order roll-off response. In contrast, the frequency response magnitude of the Otodynamics BS probe microphone displays a broad peak at 3.4 kHz as well as an undulating phase response which is attributable to the combined effect of the mechanical resonance of the microphone and tubing / resistor element.

5.3 Electroacoustic probe Thevenin parameters

The Thevenin parameters of the Otodynamics BS probe are given in figure 5.4. The B&K Type 4182 probe microphone frequency response was subtracted from the sound pressures measured in each calibration tube, rendering the open circuit pressure, $p_s$, independent of the measurement microphone.
Figure 5.4 Thevenin parameters of the Otodynamics BS probe against frequency. The upper left panel displays the open circuit pressure magnitude, $|p_s|$. The ordinate scale is given on the left in units of dB re 20 μPa/V. Open circuit pressure phase referenced to the electrical drive voltage, on the lower left panel, is shown in units of degrees and is bounded by ±200°. The upper right panel displays the series impedance magnitude, $|Z_s|$, given in units of dB re 1 M MKS acoustic ohm. The series impedance argument is shown in the lower right panel in units of degrees, ranging from -100° to +100°. Data points for all plots at 16 Hz intervals with straight line interpolation. Plotted for each parameter is the mean, corresponding to the mid curve, which is sandwiched between the 95% confidence interval as estimated from the variance-covariance matrix.

---

When specifying (driving point) impedance, argument is preferable to phase. Implicit in the notion of phase is the comparison between two signals; the signal of interest and a reference signal, which usually relate to different parts of the system, leading to the concept of a transfer characteristic. The vector of the ratio of the two signals, for increasing frequency, on the complex plane is allowed to cycle clockwise, continuously and unbounded, a characteristic relating to the concept of latency. Whereas, the argument of acoustic impedance relates to the phase of sound pressure relative to volume velocity at the same point in the system. The result is that the argument of the acoustic impedance is bounded by ±180°. The quadrant that the impedance vector occupies is physically meaningful, defining the network characteristic at the measurement point. For example, a vector in the upper right, first, quadrant indicates that the network is mass limited and passive at that measurement point. However, impedance does not relate to a system characteristic per se, because reference to an additional, secondary point in the network is not made. For example, consider a generalised case in which a two-port network is loaded on its output port by an impedance, $Z_c$. If the two-port network consists of a single series impedance element, $Z$, then the impedance at the input port is given by $Z + Z_c$. The system could contain an active element, that is $\Re{Z} < 0$, coupled to a passive $Z_c$, or vice versa, which, because of the relative magnitudes of the real parts, may lead to a purely passive characteristic in the impedance at the input port.
For complex signals, such as chirps, the power spectral density is a function of the analysis bandwidth and resolution; the bandwidth was 12.8 kHz with 801 spectral lines, i.e., a resolution of 16 Hz. Therefore, to represent $|p_s|$ independent of the signal analysis, the magnitude was rescaled so as to correspond to a single sinusoidal stimulation.

The rapid change exhibited in the phase of $p_s$, is simply a result of the wrapping effect, and is valid since the vector, $p_s$, is allowed to cycle clockwise on the complex plane continuously, sweeping through the positive real, negative imaginary, negative real and then positive imaginary axes. Such a rotating phase characteristic is in contrast to the argument of $Z$. By virtue of representing the probe in terms of its Thevenin parameters, $Z_s$ is constrained to a passive characteristic, whose argument must be bounded by $-90^\circ$ to $+90^\circ$, since $\text{Re}(Z_s)$ is non-negative.
Chapter 5: Probe assembly calibration: Results

The standardised ear canal characteristic impedance\(^3\) is 9.2 M MKS acoustic ohm, equivalent to 19.3 dB re 1 M MKS acoustic ohm (Gilman et al, 1981). At 5 kHz, \(|Z_s|\) is at a minimum of 31 dB re 1 M MKS acoustic ohm, a factor of 3.9 greater than the standard ear canal characteristic impedance. Under this condition a 10 % change in the load impedance leads to a corresponding change of approximately 8 % in the pressure developed across the load. Because the change in pressure is proportionally less than the change in load impedance, at 5 kHz the probe deviates from a constant volume velocity source.

5.4 Errors in calibration

Mismatch between the measured sound pressures and model data carries information on the validity of the calibration procedure. Statistical inference through the use of confidence intervals derived from standard errors is valid for errors which are gaussian distributed. Residual errors, the distances from the LMS estimate to the actual data point, across frequency and tubes, were normalised by calculating the standardised residual errors. The population of Guassian distributed standardised residual errors have a mean of zero and standard deviation of one. However, since the LMS minimisation was performed in two dimensional complex space, the distribution of the standardised residual errors is maximal at the origin of the complex plane and is rotationally symmetric about the origin.

Additionally, the standardised residual errors at one frequency should be independent of the standardised residual errors at neighbouring frequencies, leading to a zero correlation across frequency. The standardised residual errors within each tube were plotted to establish two points; firstly, to determine if a systematic pattern existed between neighbouring frequencies and secondly, whether the standardised residual errors are gaussian distributed in two dimensions.

\(^3\) Comparison, for the purposes of the analysis of power transfer, between the standardised ear canal characteristic impedance and the specific characteristic impedance of air, equal to 409 Kg m\(^{-2}\) s\(^{-1}\) (Keefe, 1984), is not possible because the two parameters are dimensionally different.
Qualitatively, from comparing figures 5.5 and 5.6, it can be seen that the relationship between neighbouring frequencies, i.e., the co-ordinates of consecutive dots, is less systematic post-minimisation; pre-minimisation smooth curves exist which suggest a dependence between frequencies, whereas post-minimisation a more random chaotic pattern exists reflecting less correlation across frequency.
Figure 5.5: Standardised residual errors pre-minimisation. The standardised residual error across frequency for each tube is plotted in each of the six panels. The real part of the standardised residual error is given by the abscissa, the imaginary part by the ordinate value.
Figure 5.6: Standardised residual errors post-minimisation. Format as for figure 5.5.
Table 5.1: Quantitative analysis of standardised residual errors

<table>
<thead>
<tr>
<th>Tube number</th>
<th>Mean</th>
<th></th>
<th>Standard deviation</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>pre</td>
<td>post</td>
<td>pre</td>
<td>post</td>
</tr>
<tr>
<td>1. re</td>
<td>-0.0527</td>
<td>0.0551</td>
<td>0.722</td>
<td>0.746</td>
</tr>
<tr>
<td></td>
<td>im</td>
<td>0.235</td>
<td>0.0211</td>
<td>0.849</td>
</tr>
<tr>
<td>2. re</td>
<td>-0.148</td>
<td>-0.159</td>
<td>0.554</td>
<td>0.311</td>
</tr>
<tr>
<td></td>
<td>im</td>
<td>0.114</td>
<td>-0.0802</td>
<td>0.429</td>
</tr>
<tr>
<td>3. re</td>
<td>-0.0112</td>
<td>0.382</td>
<td>0.692</td>
<td>0.447</td>
</tr>
<tr>
<td></td>
<td>im</td>
<td>-0.344</td>
<td>0.228</td>
<td>0.480</td>
</tr>
<tr>
<td>4. re</td>
<td>0.143</td>
<td>0.0912</td>
<td>0.534</td>
<td>0.637</td>
</tr>
<tr>
<td></td>
<td>im</td>
<td>0.205</td>
<td>0.187</td>
<td>0.400</td>
</tr>
<tr>
<td>5. re</td>
<td>-0.0207</td>
<td>0.0119</td>
<td>0.406</td>
<td>0.426</td>
</tr>
<tr>
<td></td>
<td>im</td>
<td>-0.0382</td>
<td>0.0582</td>
<td>0.504</td>
</tr>
<tr>
<td>6. re</td>
<td>0.0888</td>
<td>-0.0373</td>
<td>0.336</td>
<td>0.515</td>
</tr>
<tr>
<td></td>
<td>im</td>
<td>-0.172</td>
<td>0.0416</td>
<td>0.524</td>
</tr>
</tbody>
</table>

From table 5.1, quantitative analysis suggest that the minimisation procedure is inadequate in establishing standardised residual errors that are Guassian in distribution where the mean equals zero and standard deviation equals one.

5.5 Varying sound source constructions

To illustrate the effect of different probe constructions upon the Thevenin parameters, four different sound sources were temporally constructed using the same transducer. By utilising heat shrink tubing a sound source was made up with one loudspeaker, Knowles BP1712, coupled to a length of tubing, of 1 mm ID, via an acoustic resistor located at the loudspeaker port. This configuration allowed the value of the resistor and length of tubing to be changed easily while keeping the sound source assembly hermetically sealed. Resistor values 0, 68, 220 and 470 M MKS acoustic ohm and tubing lengths of 13 mm and 20 mm were used.
Figure 5.7: Effect of acoustic resistor upon Thevenin parameters. Each sound source consisted of a loudspeaker (Knowles BP1712) coupled to a port tube of length 13 mm, via an acoustic resistor. The loudspeaker was driven from a constant voltage source, via a 560 ohm resistor, with a chirp signal, -35 dBV, over a bandwidth from 0 to 12.8 kHz. The parameters in the legend refer to the acoustic resistor in units of M MKS acoustic ohm.
**Figure 5.8**: Effect of length of port tubing upon Thevenin parameters. Each sound source consisted of a loudspeaker (Knowles BP1712) coupled to a port tube via a resistor of value 2.2 k MKS acoustic ohm. The loudspeaker was driven from a constant voltage source, via a 560 ohm resistor, with a chirp signal, -35 dBV, over a bandwidth from 0 to 12.8 kHz. The parameters in the legend relate to the length of tubing in units of mm.

From figures 5.7 and 5.8 it can be seen that a change in the sound source construction affects both $p_s$ and $Z_s$. It may seem counterintuitive that a change, effectively, in the impedance of the sound source modifies the Thevenin open circuit pressure parameter, $p_s$. However, the complete sound source can be modelled as an internal sound source coupled to a network of series and shunt lumped immittance elements, manifesting in a topology of series and parallel branching. It is possible therefore that a change in the immittance elements may give rise to a change in $p_s$. 
The selection of the acoustic resistor and tube length plays a major role in the sound source Thevenin parameters. From figure 5.7, the absence of an acoustic resistor leads to $Z_s$ fluctuating against frequency from stiffness to mass limited regimes, as seen in the argument plot. The inclusion of a 470 M MKS acoustic ohm resistor effectively damps the resonances such that only one transition from stiffness to mass limited regime is observed. The modification also leads to a smoothing of $|Z_s|$ and $p_s$. Terminating the 1 mm ID tubing, of characteristic impedance 521 M MKS acoustic ohm, with a resistor of 470 M MKS acoustic ohm provides better coupling compared to the true impedance of the loudspeaker. By extending the sound source tubing by 50 %, the first impedance notch shifts downwards in frequency by 20 %, as shown in figure 5.8.

### 5.6 Making input impedance measurements

With a calibrated probe whose Thevenin parameters are shown in figure 5.4, measurements of the input impedance of unknown acoustic loads can be made, by sampling the sound pressure at the input, $p_{in}$. With reference to the potential divider equation, $Z_{IN}$ is solved for:

$$Z_{IN} = \frac{p_{IN} Z_s}{p_s - p_{IN}} \quad (5.1)$$

Measurement of $p_{in}$ is made with the microphone located in the probe assembly, whose transfer function will be compensated for.

### 5.7 Propagation of errors

Knowledge of the uncertainty in each of the Thevenin parameters allows for the errors to be propagated to the estimates of input impedances. The relative scalar error in $Z_{IN}$ is:

$$\frac{\Delta Z_{IN}}{Z_{IN}} = \frac{\Delta p_{IN}}{p_{IN}} + \frac{\Delta Z_s}{Z_s} + \frac{\Delta p_s + \Delta p_{IN}}{p_s - p_{IN}} + \frac{j R_s k a e}{Z_{IN}} \quad (5.2)$$
where $\Delta E$ is the absolute complex error in $E$, such that the relative scalar error in $E$ is given by $|\Delta E/E|$. The relative error in $p_{nw}$ is estimated at the time of measurement by noting the signal to noise ratio, while the relative error in both $p_s$ and $Z_s$ is carried through from the variance-covariance matrix estimated during the Thevenin calibration routine. The final term refers to the relative error due to the presence of the evanescent wave, which is calculated for an estimated value of $\varepsilon$.

5.8 Measured input impedance of hard walled rigidly terminated tube

Using a calibrated probe, whose Thevenin parameters are given in figure 5.4, the input impedance of a hard walled rigidly terminated tube was measured, as shown in figure 5.9.
Figure 5.9: Measured input impedance of a cylindrical hard walled tube of length 1.2 cm and radius 4.5 mm. A chirp stimulus of level -35 dBV was applied, via a 560 ohm resistor, to the loudspeaker. The microphone within the probe sampled the resulting sound pressure. Shown for comparison is the model prediction using the isothermal tube model including thermoviscous losses (Keefe, 1984). The error bars represent the ± absolute scalar error as defined in section 5.8, with $\varepsilon = -0.3$. 
Chapter 5: Probe assembly calibration: Results

The model impedance, $Z_{\text{mod}}$, displays a stiffness controlled region on the low frequency side of the resonant frequency (7234 Hz), where the magnitude decreases at a rate of 6 dB/Oct. The argument is slightly greater than -90°, indicating a small resistance that is a function of the thermoviscous losses. At frequencies above resonance, the input impedance is mass controlled with a magnitude increasing at a rate of 6 dB/Oct. and an argument just less than 90°.

At a frequency corresponding to $|Z_{\text{meas, min}}|$ the argument is not bounded by ±90° suggesting a negative real part and therefore an active component in the load. This departure from the passive case is probably due to contamination by noise because the SPL is also at a minimum leading to a decrease in the SNR.

The measured input impedance, $Z_{\text{meas}}$, deviates maximally from $Z_{\text{mod}}$ at frequencies around the $|Z_{\text{meas, min}}|$, indicative of the presence of an evanescent wave (Keefe and Benade, 1981; Brass and Locke, 1997). The model estimates the frequency of the $|Z_{\text{mod, min}}|$ to be located at 7234 Hz, whereas $|Z_{\text{meas, min}}|$ is shifted 12% higher to 8100 Hz.

Assuming a constant volume velocity drive condition, where sound pressure is directly proportional to input impedance, and with reference to figure 7 from Brass and Locke (1997), a shift upwards in frequency suggests that the evanescent mode factor is negative. At 8100 Hz, the argument of $Z_{\text{meas}}$ is zero, indicating that $Z_{\text{meas}}$ is purely resistive, whereas $Z_{\text{mod}}$ is complex, equal to $0.05 + j1.27$ M MKS acoustic ohm. Therefore, $\Im(Z_{\text{mod}}) = -Z_{ev}$, that is, $\Im(Z_{\text{mod}}) = -R_o k \alpha \epsilon$, where $R_o = 6.42$ M MKS acoustic ohms, $k = 146.6$ m⁻¹. This leads to $\epsilon = -0.3$, a figure which is equivalent to that estimated by theory in section 4.5.
5.9 Measured input impedance of Brüel & Kjær ear simulator

**Figure 5.10:** Measured input impedance of Brüel and Kjær ear stimulator Type 4157, determined by Thevenin calibrated probe. Shown for comparison are data from the Brüel and Kjær manual (1991) figure 3.3, as well as from Voss and Allen (1994) measurement on Type 4157, from figure 12. Cubic spline interpolation between data points.
Comparing figure 5.9 and figure 5.10, the minimum in the magnitude of the input impedance of an ear is approximately 26 dB greater than that measured in a hard walled rigidly terminated tube. The consequence of an increase in the minimum input impedance is that the influence of the evanescent wave upon acoustic measurements is proportionally less. The magnitude of the equivalent evanescent impedance acts as an inductive element, increasing at 6 dB/Oct, and at 10 kHz is of the order of 5 dB re 1 M MKS acoustic ohm.

The 4157 is designed to comply with the relevant standard (ANSI S3.25-1979), which specifies that the acoustical length of the standard occluded ear canal be equal to 13.9 mm. In order to make robust sound pressure measurements which are not prone to poor coupling between probe and ear simulator leading to significant leakage, the ear canal extension, DB2012, was used. The probe, including foam tip, could not be inserted fully, therefore a small volume was enclosed between the probe tip and the reference plane of the ear simulator.
5.10 Measured input impedance of adult ears

Figure 5.11: Measured input impedance of adult ear, determined by Thevenin calibrated probe. Shown for comparison are data from Keefe et al (1993) at 1/3rd octave frequencies; magnitude data from figure 1; argument data from figure 2. Cubic spline interpolation between data points.
From Keefe et al (1993), the magnitude of the adult ear canal input impedance exhibits a shallow notch compared to the measured data, B&K ear simulator and Voss and Allen (1994) data. This can be explained by that, Keefe et al (1993) data are averaged over ten subjects while the other data correspond to individual ear measurements. It is unlikely that the frequency position of the sharp notch will be constant across the ten subjects, due to both inter-subject variability and in the probe location relative to the TM (Voss and Allen, 1994). The consequence of averaging will be an ironing out of the narrow notch observed in data from individual ears, to a smooth curve as seen in Keefe et al (1993), thereby resulting in an overestimate of the minima and an underestimate of the maxima in the ear canal input impedance magnitude.

The Z\textsubscript{\text{mean}}, for the unimpaired, otological normal female adult, shown in figure 5.11, manifests a minimum magnitude of -6 dB re 1M MKS acoustic ohm at 5 kHz. The notch magnitude is of the same order as, but at a frequency lower than, the data from Voss and Allen (1994); for ten subjects the notch frequency in kHz from Voss and Allen (1994) were 6.1, 6.1, 7.8, 5.8, 6, N/A, 8.1, 7, 8.9, >10.

The action of the occluded ear canal in terms of the sound pressure in the plane of the probe tip is a transformation of the (complex) sound pressure at the TM. The phase shifting properties of such a transformation is that high frequencies are modified more than lower frequencies by virtue of the inverse relationship between frequency and wavelength. Due to the dependence of sound pressure in the calculation of input impedance, the action of the canal upon the occluded ear canal input impedance is more significant at higher frequencies.

The use of a two component model enabled the frequency shifting nature of the occluded ear canal to be investigated. The model consists of a uniform lossy transmission line, of length \(L\), terminated by a lumped impedance described by a high pass function, as in Siegel (1994). Figure 5.12 shows the effect of the parameter \(L\) upon the occluded ear canal input impedance. Qualitatively, the frequency location of the magnitude notch decreases for increasing \(L\).
Figure 5.12: Effect of the occluded ear canal length upon the ear input impedance. Parameter is length of occluded ear canal in cm.
From the finding that the notch frequency is inversely proportional to the occluded ear canal length, it is argued that the downward shift of the notch frequency seen in the input impedance of adult ears (figure 5.11) compared to other data (Voss and Allen, 1994) is a consequence of the insertion depth of the probe being less, and therefore the occluded ear canal length being greater, in the former measurement.
Chapter 6: Ear canal acoustic measurements

6.1 Introduction

A description of the method of continuous tonal ear canal sound pressure measurement recorded at various stimulus levels will be made, followed by the presentation of experimental results. Information on the subjects who participated in the study will be given, together with a discussion on the influence of SOAEs upon ear canal acoustic measurements. At relevant points in the presentation of experimental data, the analysis of the acoustic response of the occluded ear canal in terms of system functions is developed in order to quantify and model the observed microstructure and nonlinearity.

6.2 Nature and influence of Spontaneous OAEs

Narrow band spectral analysis of the sound pressure in the sealed ear canal under no stimulation reveals, in certain ears, a series of peaks of energy with frequency. The energy peaks which are interpreted as spontaneous otoacoustic emissions (SOAEs) are characteristically narrow bandlimited; -3 dB bandwidth typically 1.2-4.7 Hz (Kemp, 1981).

When using a pressure microphone, the measurement position must be within, rather than exterior to, the ear canal due to the almost 100% reflection of the lateral travelling sound pressure wave at the concha/canal interface. Sealing the ear canal with an electroacoustic probe assembly effectively terminates the ear canal with a high impedance relative to the canal characteristic impedance, ensuring a high sound pressure is developed across the probe tip and achieving a relatively high signal to noise ratio.

Many studies model SOAEs as resulting from noise-perturbed limit-cycle oscillations (Bialek and Wit, 1984; Van Dijk and Wit, 1990; Talmadge et al, 1990). By characterising a system in its phase space, i.e., a space representing all possible states of
a system for all time\(^1\), a categorisation can be made depending upon the curve traced out, or trajectory followed, as the system state changes with time and evolves. The class of trajectories relating to dissipative systems is generically known as an attractor, and can be; a single point in phase space; a closed curve indicative of periodic behaviour (limit-cycle); or a fractal (strange attractor), representative of chaotic systems. The Van der Pol oscillator displays a limit cycle behaviour, provides a theoretical framework and accounts for many of the characteristics of SOAEs (Tubis and Talmadge, 1998).

Both experimentally (Kemp, 1979a; Wilson, 1980b) and theoretically (Van Dijk and Wit, 1990; Tubis and Talmadge, 1998) it has been shown that under an externally applied sinusoidal stimulus an SOAE manifests either suppression or phase locking, dependent upon the intensity and frequency of the external tone; the SOAE either oscillates at a frequency different from the stimulus and is suppressed, referred to as non-entrainment, or locks to the stimulus frequency, a condition known as entrainment.

Contrary to the definition of spontaneous otoacoustic emissions, in that they are non-evoked OAEs, a common method exploits the fact that SOAEs can be synchronised to an externally applied stimulus (Otodynamics ILO88 OAE analyser in SOAE Search Mode). Prominent spectral components of the synchronised response elicited from a click stimulus up to 80 ms post stimulus constitute SOAEs. Wable and Collet (1994) questioned whether a synchronised response constitutes, or can be attributed to, a purely non-evoked SOAE response. By comparing the SOAE response recorded using the evoked, synchronised method to the non-evoked method, it was concluded that the two responses were not identical in spectral detail; the evoked, synchronised response manifested energy at frequencies not present in the non-evoked response, and vice-versa. The probe used to elicit the evoked, synchronised response consisted of a microphone and loudspeaker, whereas, the non-evoked method employed a probe comprised of a single microphone of different dimensions. The finding from chapter 5, that probe construction, especially transducer port tubing dimensions, strongly

\(^1\) An \(n\) degree of freedom system leads to a \(2n\) dimensional phase space. For example, a phase space of \(6N\) dimensions is required to represent a system of \(N\) point particles because there are three positional co-ordinates and three momentum co-ordinates. A one degree of freedom oscillator results in a two dimensional phase space with axes of force against speed.
influences the Thevenin source acoustic impedance \( Z_s \) together with observations that OAEs are a function of \( Z_s \), casts doubt on both the appropriateness in the comparison of data from the two methods, and the validity of conclusions made, in Wable and Collet (1994).

Wit _el al_ (1981) found that in a particular frequency range the synchronised response matched that of the non-evoked response, while for another frequency range the synchronised method introduced energy at frequencies not present in the non-evoked case. Ruggero _et al_ (1983) while investigating one prominent SOAE found no difference in the spectra from the two methods.

### 6.3 Subjects

Four adult subjects, one male (PS), three females (SS, KH, JE), aged between 18 and 27 years, participated in the study. Pure tone auditory thresholds were measured between 0.25 and 8 kHz in octave steps on a GSI 10 audiometer, calibrated to BSI/ISO standards in order to screen for normal hearing. Subjects’ thresholds recorded in both left and right ears fell within -10 to +20 dB HL, defined as the normal audiometric range.

Spontaneous OAEs, recorded using the Otodynamics ILO88 OAE analyser in SOAE Search Mode, i.e., a synchronous method to measure SOAEs, between 0 and 6.4 kHz, are shown in figure 6.1, and are indicated on subsequent data where germane.
Chapter 6: Ear canal acoustic measurements

Figure 6.1: Narrowband spectral analysis of ear canal sound pressure against frequency for each subject. Recorded using the ILO88 in SOAE search mode. Spectral bandwidth equal to 12 Hz. Bottom panel displays the energy reflectance at the probe tip in terms of retrograde, or backward, energy propagation, i.e., ratio of forward travelling ear canal energy to backward propagating ear canal energy. The curve represents the energy reflectance of a lossless uniform transmission line of radius equal to the standardised ear canal radius, 3.75 mm (ANSI S3.25-1979) terminated by the probe source impedance, $Z_s$, given in figure 5.4. Since the line is lossless the energy reflectance is independent of line length.
It can be seen from figure 6.1 that the SOAE against frequency appears to be unique for
the ear under test, displaying no intersubject correlation. The presence, and frequency
location, of an SOAE may, however, be a function of the acoustical characteristics of
the recording probe.

6.4 Influence of the probe upon SOAEs

By using an analogue model of the auditory periphery coupled to a probe of a specified
source impedance, $Z_s$, Zwicker (1990) observed a small shift, of the order of 0.5 %, in
the frequency location of an SOAE between an inductive and capacitive $Z_s$
characteristic.

More importantly in terms of the influence of the acoustical characteristics of the
measuring device upon the measurement of SOAEs, Allen et al (1995) suggests that a
correlation exists between the reflectance magnitude of the probe and the prevalence of
an SOAE; the higher the reflectance magnitude the more prevalent the SOAEs. Probe
reflectance magnitude decreases as $|Z_s|$ tends towards the characteristic impedance of
the ear canal. Referring to the Thevenin parameters of the probe (figure 5.4), $|Z_s|$ tends
towards the characteristic impedance of a typical ear canal, approximately 20 dB re 1 M
MKS acoustic ohm, at around 4.75 kHz, a region where there are no measurable SOAEs
(figure 6.1). The energy reflectance, i.e., the square of the pressure reflectance
magnitude is given in the bottom panel of figure 6.1 and displays a minimum around 5
kHz.

The relationship between SOAEs and probe reflectance can be revealed qualitatively by
the following argument. Model the SOAE as a perfect sound pressure source, $P_{SOAE}$,
coupled to a series source impedance, $Z_{SOAE}$, equivalent to the forward ear canal input
impedance at the TM. The SOAE Thevenin equivalent source is loaded by the occluded
ear canal, whose characteristic impedance is $Z_{ec}$, and is terminated by the
electroacoustic probe assembly source impedance, $Z_s$. Under the condition of the ear
canal perfectly terminated by $Z_s$, the probe tip sound pressure, $|P_{ec}^{match}|$, is that sound
pressure launched onto the occluded ear canal at the TM and governed by $Z_{ec}^o$ relative to $Z_{SOAE}$. Since $Z_{ec}^o$ is dominated by its real part, whose standardised value equals 9.2 M MKS acoustic ohms (Gilman et al, 1981), the matched condition occurs when $Z_s$ is resistive.

However, when $Z_s$ is not equal to $Z_{ec}^o$, $|p_{ec}|$ can vary from the matched case for two reasons. Firstly, the load input impedance is modified, resulting in a change in the absolute sound pressure launched onto the ear canal, the level of which is again determined by the ratio of the load input impedance to $Z_{SOAE}$. And secondly, the change in $Z_s$ effectively modifies the pressure reflectance leading to a change of sound pressure but bounded by zero and $2 |p_{ec}^{match}|$, governed by $Z_s$ relative to $Z_{ec}^o$.

Consider for the purposes of a qualitative treatment, that the equivalent Thevenin source working into a load impedance equal to $Z_{ec}^o$ acts as an ideal power source. Consequently, changes in the sound pressure launched onto the canal are small relative to changes arising in the load input impedance. For example, a change of 20 dB (x10) in $Z_{SOAE}$ leads to a 5.2 dB (x1.8) change in $|p_{ec}|$. However, caution must be exercised in such a coupling change. When $|Z_s| >> Z_{ec}^o$ the load input impedance magnitude undulates with frequency about $Z_{ec}^o$, such that at the quarter wavelength frequency the load input impedance is at a significant minimum. The quarter wavelength frequency for the standardised occluded ear of length 1.26 cm is roughly 6.9 kHz. For longer occluded ear canal lengths the quarter wavelength is proportionally less. At the quarter wavelength frequency, the input impedance is at a minimum and consequently the sound pressure launched onto the ear canal is also at a minimum. It is postulated that such an effect, although to a lesser degree since $|Z_s|$ is falling with frequency, is a contributory factor in the lack of SOAEs evident about 5 kHz.
Chapter 6: Ear canal acoustic measurements

The relationship between $p_{ec}$, $Z_S$ and $R$ is given below, valid for one dimensional motion along a uniform transmission line.

\[ p_{ec} = p_{ec}^- (1 + R^a) \]  \hspace{1cm} (6.1)

\[ R^a \equiv \frac{p_{ec}^*}{p_{ec}} = \frac{Z_S - Z_{SOAE}}{Z_S + Z_{SOAE}} \]  \hspace{1cm} (6.2)

From the two equations above it is seen that $R$ is complex when $Im\{Z_S\} \neq 0$, and $R$ is bounded by a circle of radius equal to one, therefore $0 \leq |p_{ec}| \leq 2|p_{ec}^{\text{match}}|$ and finally, $|p_{ec}|$ is maximal when $|Z_S| >> |Z_{SOAE}|$.

The data in figure 6.1 provides evidence to support the finding in Allen et al (1995) of a dependence of $Z_S$ and the effect of the quarter wavelength frequency upon the measurement of SOAEs in human adults, highlighting the desirability of a (very) high source impedance for analysis frequencies below 6 kHz.

It was considered desirable to minimise the effect of SOAEs upon the measurement of CTASP microstructure due to the, potentially confounding nature of the two interacting auditory phenomena. The ear under test, left or right, was chosen as that ear which manifested the less SOAEs in the two octave band region centred upon 1.5 kHz, determined simply by inspection.

6.5 Ear canal acoustic measurement system

Subjects were seated in an upright position on a comfortable upholstered chair, in a sound treated room. To reduce background noise, the computer base unit and lock-in amplifier, which both contained fans, were located in the anteroom of the sound treated room.
Chapter 6: Ear canal acoustic measurements

Figure 6.2: Schematic of the experimental set-up. The probe microphone was of electret condenser technology with an internal F.E.T. preamplifier which acted as a buffer. A d.c. supply (±9 V) was required to power the F.E.T. A 560 ohm resistor, in series with the 50 ohm output impedance of the oscillator, ensured that the passive attenuator was correctly loaded on one side. With the attenuator loaded on the other side by the loudspeaker an adequate accuracy in attenuation was achieved. The function of the attenuator was to enable a high level sound pressure to be developed in the ear canal without restricting the amplitude range of the signal generator. Changes in the stimulus amplitude were achieved by varying the signal generator amplitude under PC-control. During the course of a single frequency sweep the voltage applied to the loudspeaker was constant; hence, the acoustic stimulus was generated under a constant applied voltage drive condition. Stimulus frequency was stepped at discrete values, typically at 5 Hz intervals, but occasionally 10 Hz intervals at higher frequency regions, in order to characterise the microstructure with adequate resolution.

The EG&G DSP dual channel lock-in amplifier Model 7220 enables the recovery of a sinusoidal signal that is embedded in a complex signal. A lock-in amplifier is a synchronous amplifier in which a reference, in this case internal, free-running oscillator synchronises to an input signal. The analysed signal is in the form of two DC outputs,
one DC output proportional to the in-phase component of the input signal, the other proportional to the quadrature (90°) component of the input signal, a configuration known as quadrature phase. An AC signal superimposed upon the DC outputs is representative of bandlimited noise of a bandwidth equal to that of the analysis resolution and centred at the reference frequency. Since the vector input signal is completely described by the two orthogonal components, a lock-in amplifier is sensitive to the magnitude and phase of the input signal.

The basic operation of a lock-in amplifier is as follows; the reference signal is converted to a square wave, which is then multiplied with the input signal, the product is then low pass filtered, with an associated filter time constant $T_c$ and skirt slope, to give the DC output. The value of $T_c$ determines the amount of averaging; the longer the $T_c$, the smoother the averaging, and hence the better SNR. The value of $T_c$ is reciprocally related to the bandwidth of the signal analysis, so that $T_c=500$ ms leads to an effective analysis bandwidth of 2 Hz, centred upon the reference frequency. The output does not follow changes in input instantaneously, but rather its response is exponentially related, such that when the input changes by a factor $k$, a settling time equal to $5 T_c$ is required for the output to reach its true value to within 5%. The choice of $T_c$ is therefore a compromise between noise rejection and the frequency sweep rate.

Typically, for stimuli levels of about 60 dB SPL $T_c=50$ ms was chosen, while for stimuli levels around 20 dB SPL, $T_c=500$ ms, i.e., a tenfold decrease in bandwidth, proved expedient because the signal was proportionally more contaminated by noise.

The 7220 was interfaced to an IBM compatible PC via an RS232 interface permitting instructions to be sent to, and data read from, the lock-in amplifier. Additionally, the 7220 internal oscillator allows for computer controlled generation of a sinusoidal signal which is used as the internal reference signal, as well as the external, ‘Osc. out’, signal.
6.6 Continuous tonal aural data

Figure 6.3: Sound pressure measured in the occluded ear canal, $p_{oc} [\omega, A]$, of subject SS at two stimulus levels; thick line: $A = 60\,$dB SPL. thin line: $A = 20\,$dB SPL. The upper panel displays $|p_{oc} [\omega, A]|$, the lower panel $p_{oc} [\omega, A]$ phase relative to the electrical drive signal. Note that the curve representing the high level stimulation is read from the left ordinate while the right ordinate corresponds to the low level stimulus data. Datum at every 5 Hz, with cubic spline interpolation between data points. Vertical dotted line indicates frequency location of spontaneous otoacoustic emission, data from figure 6.1. A peak in the narrow band spectrum of the synchronised response was labelled a spontaneous emission for signal to noise ratio greater than 6 dB. Uncertainty in frequency of SOAE is ± 6 Hz.
Calibration of the stimulus level, that is, adjustment of the \textit{in situ} stimulus level to some predetermined magnitude, was made at the high level to a resolution of $\pm 1$ dB SPL, at the mid frequency of the sweep only. As no compensation was made to ensure a constant sound pressure level at other frequencies, $|p_{ec}|$ during high level stimulation may not remain perfectly horizontal, but may rather exhibit a small gradual variation, reflecting a change in the gross response of the ear with frequency, as well as a deviation from a flat magnitude frequency response of the stimulus delivery and recording apparatus.

It may appear counter-intuitive to label the low level magnitude curve '20 dB SPL' since it patently manifests an undulation around 20 dB SPL. To clarify, the parameter $A$ refers to the \textit{stimulus} (or input) sound pressure level in dB SPL whereas, the low level magnitude curve in figure 6.3 describes the \textit{response} (or output) of the occluded ear canal under such a stimulus.

Two previous studies (Zwicker and Schloth, 1984; Shera and Zweig, 1993) adopted the absolute auditory threshold as the decibel reference level, i.e., dB SL. However, it was not made explicit whether the threshold was established at one or multiple frequencies in the measurement frequency range. It is assumed here that the determination of threshold at multiple frequencies appears unlikely since both studies used a stimulus continuously swept in frequency, and in the Shera and Zweig (1993) paper no mention was made relating to threshold methodology. This issue is important in view of Elliot’s (1958) finding, in that absolute auditory threshold often fluctuates against frequency with an amplitude of the order of 12 dB and a periodicity of 100 Hz in the mid frequency range, leading to the observation that a decibel scale referenced to sensation level is a function of the reference frequency. In view of this confounding effect, as well as the desire to express the acoustical response of the auditory periphery in terms of purely physical parameters, the current study used the conventional dB SPL scale.
Figure 6.4: Sound pressure measured in the Brüel and Kjær ear simulator Type 4157, at two stimulus levels; 60 dB SPL and 20 dB SPL. Lock-in amplifier and instrumentation settings as for figure 6.3.

Figure 6.3 displays a level dependent characteristic in continuous tonal aural sound pressure (CTASP) data which at a stimulus intensity of 60 dB SPL manifests a smooth response with frequency, while at stimulus intensity 40 dB lower uncovers an undulating pattern. For a stimulation level equal to 20 dB SPL, the series of maxima and minima of the undulation appear to be regularly spaced, manifesting a spectral periodicity: thus, the undulating pattern is termed microstructure. In contrast, figure 6.4 shows that sound pressures measured in an ear simulator result in the low level sweep approximately superimposing upon the appropriately scaled, high level sweep, without an undulation of cogent periodicity.
The amplitude of microstructure, peak to valley distance, for 20 dB SPL stimulation is roughly 3 dB. Additionally, the microstructure displays a spectral periodicity, interpeak frequency $\Delta f$, approximately equal to 80 Hz, equivalent to a frequency spacing of $\frac{1}{12^{th}}$ of an octave, or $\frac{\Delta f}{f} = \frac{1}{17}$. Referring to the phase plot, at 20 dB SPL the phase varies by up to $\pm 0.2$ rad ($\pm 12^\circ$) from the high level curve.
Figure 6.5: CTASP data for four subjects. Lock-in amplifier and instrumentation settings as for figure 6.3. Thick line corresponds to $A=60\,\text{dB SPL}$, thin line $A=20\,\text{dB SPL}$. Vertical dotted line indicates frequency location of spontaneous otoacoustic emission data, from figure 6.1.
Chapter 6: Ear canal acoustic measurements

The best frequency region, i.e., the region in which the CTASP microstructure is most prominent, for each of the four subjects is shown in figure 6.5. The top left panel focuses on one cycle in the microstructure in the data shown in figure 6.3. The top right panel shows 2½ cycles of CTASP microstructure between 900 and 1100 Hz measured in a female subject (KH), the bottom left panel shows almost three cycles of microstructure, measured in the male subject (PS), while 7½ cycles are shown in the bottom right panel for the third female subject (JE).

With reference to the influence of spontaneous otoacoustic emissions upon the ear canal acoustic measurements and comparing figures 6.1 and 6.5, CTASP microstructure is independent of the presence of a SOAE. Narrow band analysis of the synchronised ear canal response revealed SOAEs, defined as spectral peaks greater than 6 dB above noise, within the CTASP measurement bandwidth for subjects SS and KH as indicated by vertical dotted lines, but not for subjects PS and JE.

Qualitatively, the CTASP undulation evinces periodic and amplitude characteristics commensurate with physical data from previous reports of auditory microstructure (Wilson, 1980a, 1980b; Kemp, 1979a; Kemp and Chum, 1980; Zwicker and Schloth, 1984; Shera and Zweig, 1993; Allen et al, 1995). It is argued here that the presence of CTASP microstructure in real ear data (figures 6.3 and 6.5) and its absence in ear simulator data (figure 6.4) is neither a function of the Thevenin parameters derived during the calibration procedure, nor an artefact of the measurement technique, be it of a random origin or deviation from linearity. The data confirm the existence of CTASP microstructure as a measurable auditory signal relating to the mechanics of the auditory periphery, which, due to the mechanical coupling afforded by the middle ear complex, is thought to be modified by the nonlinear response of the cochlear.

6.7 CTASP spectral periodicity

Applying the Fourier transform to CTASP data allows for quantitative analysis of the spectral periodicity, in that the peak of the amplitude of the transformed signal corresponds to the most prominent periodic component. Transforming a waveform, i.e., a time domain description of a signal, to a spectrum in the frequency domain, reveals
energy at underlying, constituent frequencies within the waveform. Similarly, a transformation of a spectrum reveals energies at particular periods that combine to characterise the spectrum. The operation transforms a spectrum as a function of a frequency variable, $f$ in Hz, to a waveform of a temporal variable, $t$, in s. Additionally, the reciprocal of $t$, is given by $f$, in Hz.

In linear units of pressure, i.e., pascals, indicated by the lower case $a$'s, the normalised low stimulus level CTASP data is defined as:

$$p^a_{e}[\omega, a_L, a_H] = \frac{a_H}{a_L} \frac{P_{ec}(\omega, a_L)}{P_{ec}(\omega, a_H)}$$

The Fourier transform of the $p^a_{ec}[\omega, a_L, a_H]$ vector must be computed, rather than simply the vector amplitude, since a vector of uniform amplitude can possess a periodicity in the in-phase and quadrature components related to the angular velocity of the vector. For example, a rotating vector of uniform amplitude, $e^{\omega t}$ where $\omega = 2\pi f$, manifests a periodicity equal to $1/f$.

**Figure 6.6:** Quantitative analysis of CTASP microstructure periodicity. a-d) Amplitude of the Fourier transform of CTASP data, in Pa, from subject JE given in figure 6.5. Represented as a function of four variables; a) absolute spectral periodicity, $t_a$, b) absolute spectral frequency, $f_a$, c) dimensionless spectral period and d) dimensionless spectral frequency. e) real part of autocorrelation function against lag frequency in Hz.
Chapter 6: Ear canal acoustic measurements

JE

1.0 to 1.4 kHz
1.4 to 1.8 kHz
1.0 to 1.8 kHz

0.4
0.3
0.2
0.1
0.0
0.0 0.05 0.10 0.15 0.20
0.6
0.5
0.4
0.3
0.2
0.1
0.0
10 20 30 40 50
0.1
0.0 0.01 0.02 0.03 0.04
0.6
0.5
0.4
0.3
0.2
0.1
0.0
1 0 100 1000

Spectral period, \( t_s \) (s)
Spectral frequency, \( f_s \) (Hz)

Dimensionless spectral period
Dimensionless spectral frequency

Normalised correlation
real part

\( \tau \) (Hz)

page 103
Figure 6.6a-d shows the amplitude of the Fourier transform applied to $p_{ec}^n(\omega, a_L, a_H)$ from subject JE, as shown in figure 6.5. The measurement bandwidth is almost an octave, encompassing 7½ cycles of microstructure. As discussed in section 2.X, auditory microstructure appears periodic over a small number of cycles, but when the bandwidth is greater, the periodicity is observed to be a function of frequency; spectral periodicity is approximately proportional to stimulus frequency.

Quantitatively, the spectral periodicity observed from previous studies of CTASP microstructure approximately equals 0.4 Bark (Zwicker, 1989), exhibiting a normalised interpeak frequency spacing, i.e., $\Delta f / f$, of roughly 1/15 (Zweig and Shera, 1995), equating to a figure slightly greater than 1/12th of an octave, or a semitone.

Such a logarithmic dependency is seen in figure 6.6. The Fourier transform of the spectrum, expressed as a function of linear frequency, is plotted for two frequency regions (1 to 1.4 kHz and 1.4 to 1.8 kHz). The peak in the amplitude distributions differ, as displayed in panels a) and b); the most prominent spectral frequency of microstructure occurs at 65 Hz for the lower frequency region and rises to 100 Hz for the upper frequency region. Using $\Delta f / f \approx 1/15$, the two figures compute to 80 Hz and 107 Hz, respectively. Transforming the data to a logarithmic frequency axis, the amplitude of the Fourier transform peaks at one distinct value for the complete frequency range (1 - 1.8 kHz), as shown in panel c) and d). The coordinates of panel d) correspond to those in figure 2a) in Zweig and Shera (1995).

The autocorrelation function, $r_{xx}(\tau)$, provides an alternative method to determine periodicity in $x(\omega)$. Performing an autocorrelation can be thought of as sliding the function $x(\omega)$ in $\omega$, i.e., $x(\omega + \Delta\omega)$ over $x(\omega)$ fixed in $\omega$ and assessing the degree of similarity between the two functions; the greater the similarity the higher $r_{xx}(\omega)$ which is bounded by ±1. Therefore, the autocorrelation function is intimately linked with the notion of a periodic function, defined as $x(\omega) = x(\omega + \Delta\omega)$, such that the period is equal to $\Delta\omega$. Note that, when applied to time sequences the autocorrelation
is a function of time lag, whereas since there is no restriction upon the independent variable of $x$, the above discussion considers the autocorrelation as an interval in $\omega$.

Noting the similarity between the manipulations required to compute $r_{xx}(\tau)$ and the process of convolution, implementation of the autocorrelation function is most elegant when performed as a product of two Fourier transforms, such that:-

$$r_{xx}(\tau) \leftrightarrow F\{x(\omega)\}.F\{x(\omega)\}^*$$  \hspace{1cm} (6.4)

where $^*$ indicates the complex conjugate, leading to:-

$$r_{xx}(\tau) \leftrightarrow |F\{x(\omega)\}|^2$$  \hspace{1cm} (6.5)

where $\leftrightarrow$ implies a transform pair and $F\{\cdot\}$ indicates a Fourier transform (Press et al, 1994). However, although the application of the autocorrelation function to real valued functions is established, its utility when applied to complex valued functions is diminished since caution must be exercised in the information carried in $r_{xx}(\tau)$. To illustrate this claim consider the simple case of the autocorrelation of a vector rotating on the complex plane with unity magnitude.

Defining the autocorrelation function as:-

$$r_{xx}(\tau) = \frac{\int_{a+\tau}^{a} x(t).x(t+\tau) \, dt}{\int_{a}^{a+\tau} x(t).x(t) \, dt}$$  \hspace{1cm} (6.6)

and using Euler’s complex exponential function to describe the vector:-

$$x(t) = \exp(i\omega t)$$  \hspace{1cm} (6.7)

leads to:-

$$r_{xx}(\tau) = \exp(i\omega \tau)$$  \hspace{1cm} (6.8)
It is obvious that the information pertaining to periodicity, unlike the preceding method, is not carried in the magnitude of $r_{xx}(\tau)$, but rather the real and imaginary parts which are cosine and sine functions respectively. The quadrature components peak at intervals equal to multiples of the fundamental period, unlike the preceding method where the fundamental period dominates.

The autocorrelation of CTASP data is shown in figure 6.6e which displays a series of maxima and minima against lag frequency, which lie within ±1. However, the maxima values do not equal unity because the CTASP data deviates from a regular periodic function and is contaminated by noise.

It is concluded that the magnitude of the Fourier transform of the spectrum provides a more convenient quantitative analysis of microstructure periodicity over the autocorrelation function.

6.8 Measurements of CTASP at different stimulus levels

So far we have considered stimulating the ear at two levels, namely 60 and 20 dB SPL. Figure 6.7a displays CTASP swept at stimulus levels between 5 to 60 dB SPL at 5 dB SPL steps, across a narrow frequency region for subject KH. A spontaneous OAE was known to fall within the frequency region.
Figure 6.7: CTASP data at various stimulus levels.  

a) Left column for subject KH. 
Stimulus level varied from 5 to 60 dB SPL in 5 dB steps. The lower the stimulus level 
the more evident the microstructure in the response. Vertical dotted line indicates 
frequency of known spontaneous OAE.  
b) Right column data from subject SS at 5 
stimulus levels. The upper panel displays the normalised magnitude ear canal sound 
pressure; the lower panel the phase of $p_{\alpha} \left[ \omega, A_i \right]$. The parameters in the legend refer 
to the stimulus level in dB SPL.

In order to display $|p_{\alpha} \left[ \omega, A_i \right]|$ for many values of $A$, the magnitude data is normalised 
by the corresponding stimulus level, i.e., $|p_{\alpha} \left[ \omega, A_i \right]| - A_i$. In figure 6.7, CTASP data at 
five stimulus levels is shown; 60, 40, 30, 20 and 10 dB SPL for subject SS.
It is seen from figure 6.7, that the microstructure amplitude and phase response is inversely proportional to the stimulus level; as the stimulus level is reduced the oscillatory component becomes more prominent. However, ignoring the fluctuations associated with random noise which is most prominent in $P_{ec} [\omega, 10]$ data, as shown in figure 6.7b, the microstructure corresponding to a stimulation of $A = 20$ dB SPL is closer in amplitude to $A = 10$ dB SPL than for $A = 30$ dB SPL. Such a finding suggests that at the lowest stimulus levels CTASP data tends towards linearity since approximate homogeneity is observed. Zwicker and Schloth (1984) and Shera and Zweig (1993) report linear behaviour in ear canal acoustics at low stimulus levels.

A level independent response at stimulus levels below 20 dB SPL is observed in other auditory measures. For example, inner hair cell d.c. intracellular potential increases in direct proportion to stimulus level over comparably low levels (Russell and Sellick, 1978), synchronisation of auditory nerve fibre firing times increase linearly near absolute threshold (Littlefield, 1973) and direct measurements of basilar membrane velocity reveals a linear response at low intensities (Robles et al, 1986).

Identifying two regimes of linear behaviour in auditory measures, i.e., above 60 dB SPL and below 20 dB SPL, allows simple linear network analysis to be applied to data elicited within the linear regimes.

6.9 Network analysis of CTASP data

The following discussion borrows from conventional electrical network analysis, the theory of which is covered in Ruston and Bordogna (1966). Under harmonic stimulation, in the steady state, the measurement of $P_{ec} [\omega, A]$ can be considered the output signal resulting from the an input signal, $P_s [\omega, A]$, applied to a network, $N_{ec}$. For our purposes, $P_s [\omega, A]$ rather than the electrical voltage applied to the sound source, corresponds to the input signal since the open circuit pressure source carries information on the applied voltage and is therefore appropriately scaled by the stimulus amplitude, $A$. Characterisation of the network response, in the frequency domain, is
facilitated through the network transfer function\(^2\), \(H_{ne}(\omega, A) = \frac{p_{ne}(\omega, A)}{p_s(\omega, A)}\), which is dimensionless because both \(p_{ne}(\omega, A)\) and \(p_s(\omega, A)\) have dimensions equivalent to sound pressure, i.e., \(ML^{-1}T^{-2}\). Assuming linearity and time-invariance (hence the term LTI) in the response of \(N_{ne}\), \(H_{ne}(\omega, A)\) completely describes the network, independent of input and output signals. Expressed in its exponential form, 
\[H_{ne}(j\omega, A) = e^{\gamma[j\omega]} = e^{\alpha[\omega] + j\beta[\omega]},\]
where \(\gamma[\omega]\) is the exponential function, \(\alpha[\omega]\) is the attenuation function measured in nepers and \(\beta[\omega]\) is denoted the phase function in radians.

Consider the case in which \(H_{ne}(\omega, A)\) is the sum of two independent networks, one level independent, \(N_{Li}\), the other level dependent, \(N_{LD}\). Since the response of \(N_{Li}\) is irrespective of the stimulus amplitude, i.e., linear, we can assign a network transfer function \(H_{ne}^{L_i}(\omega)\) to \(N_{Li}\). Conversely, the response of \(N_{LD}\) is a function of the stimulus amplitude and therefore the transfer function, \(H_{ne}^{LD}(\omega, A)\), is a function of stimulus level, \(A\), in order to make explicit the level dependency. The network topology for such a case is shown in figure 6.8.

\(^2\) Often termed system function.
Figure 6.8: Block diagram relating network parameters applicable for the analysis of CTASP data. a) network parameters to describe the total ear canal response, b) the separation of the total network into two independent networks, one level dependent, the other level independent. Note that this figure does not relate to a physical realisation of a network, rather the relationship between the network parameters.

The total ear canal sound pressure $p_{ec}[^{\omega, A}]$ is made up of two sound pressures; $p_{ec}^{LU}[^{\omega, A}]$ resulting from the transformation $H_{ec}^{LU}[^{\omega}]$ on $p_{S}[^{\omega, A}]$, and $p_{ec}^{LD}[^{\omega, A}]$ the transformation of $H_{ec}^{LD}[^{\omega, A}]$ on $p_{S}[^{\omega, A}]$. From CTASP data recorded at a moderate stimulus level, ~60 dB SPL, $H_{ec}[^{\omega, A}]$ is primarily a function of $H_{ec}^{LU}[^{\omega}]$ since $p_{ec}^{LD}/p_{ec}^{LU} \approx 0$. It is therefore taken that $H_{ec}^{LU}[^{\omega}]$ is a function of the gross response of occluded ear canal, middle ear and cochlea, at moderate stimulus levels. At 20 dB SPL $p_{ec}^{LD}/p_{ec}^{LU} > 0$, rendering $H_{ec}[^{\omega, A}]$ to be a function of both networks. Since it is argued (section 6.X) that the level dependent signal is of cochlear origin, the network $N_{LD}$ then describes the response of a cochlear dependent transfer function as observed in the ear canal.
Figure 6.9: Transfer function derived from CTASP data shown in figure 6.3. Plotted is the transfer function relating to the total response $H_{\alpha}[\omega, A]$, which is the sum of two transfer functions $H_{\alpha}^{LD}[\omega, A]$ and $H_{\alpha}^{LI}[\omega]$. Top panel plots the attenuation function $\alpha[\omega]$ in nepers and the bottom panel the phase function $\beta[\omega]$ in radians. Subject SS.

It is noted that $H_{\alpha}[\omega]$ carries information on the source impedance $Z_S$, since rearrangement of the potential divider equation (equation 2.1) gives:

$$H_{\alpha} = \frac{P_{\alpha}}{P_S} = \frac{Z_{cc}}{Z_{cc} + Z_S}$$

(6.9)

such that both $\alpha$ and $\beta$ are a function of the interaction of the source and load.
Normalising $p_{ec}$ by $p_s$ removes the delay due to the stimulus delivery, i.e., the time taken from the voltage applied to the launching of an acoustic wave. From figure 5.4 the delay is given by the negative slope of the phase. It is argued in Shera and Zweig (1993) that removal of the stimulus delivery delay, results in a minimum-phase function.

The advantage of the $H_{ec}[\omega]$ representation over $p_{ec}$ is the removal of the Thevenin parameter $p_s$ from the data, reducing the influence of the probe upon the measurement.

Additionally, any CTASP, i.e., $p_{ec}$, measurement is a function of both Thevenin probe acoustical parameters, the significance of which was investigated by Zwicker (1990b), in which the data clearly shows a dependency upon the magnitude of the probe impedance, such that the effect of different probe characteristics shift the microstructure in frequency. The influence of the probe impedance upon the phase microstructure was, however, not given. Caution must be exercised in the analysis of CTASP signals for the purposes of inferring characteristics of the mechanics of auditory periphery, due to the possible role of the probe acoustical characteristics.

6.10 Nature of the subtracted sound pressure

Figure 6.10 is a vector representation of $p_{ec}[\omega,A]$, illustrating the behaviour of the two components with increasing frequency, at a stimulus level of 20 dB SPL, as in Kemp and Chum (1980). Each panel displays sound pressure in linear units, with the abscissa representing the in-phase component and the ordinate the quadrature component. In the figure, the $p'_{ec}[\omega,60]$ vector represents the appropriately scaled $p_{ec}[\omega,60]$ vector and can be thought of as the true stimulus applied to the ear, which is almost constant with increasing frequency.

Under the constraint of linearity, the $p'_{ec}[\omega,A_H]$ vector for any $A_H$, must superimpose upon $p_{ec}[\omega,A_L]$ for all $\omega$. However, the low level ear canal sound pressure vector, $p_{ec}[\omega,20]$, rotates about a point coinciding with the co-ordinate of $p'_{ec}[\omega,60]$. In order
to model the pattern with one additional vector, $\mathbf{p}'_{\text{ec}}[\omega, 60]$ must be modified by a rapidly rotating vector, $\mathbf{p}_{\text{em}}[\omega, 20]$, representing the level dependency, i.e., nonlinearity in the response at $A_L$. Under such a scaled subtracted paradigm the vector $\mathbf{p}_{\text{em}}$ is termed the stimulus frequency OAE (SFOAE) (Kemp and Chum, 1980). For the purposes of generality, in figure 6.10 it is taken that $|\mathbf{p}_{\text{em}}[\omega, 20]|$ exhibits a unspecified characteristic with increasing frequency as illustrated by the irregular shape the $\mathbf{p}_{\text{em}}[\omega, 20]$ vector sweeps out.
Figure 6.10: Vector representation, from Kemp and Chum (1980), of $p_{ec}[^{\omega}, A]$ at two stimulus levels; $A_H=60$ dB SPL, $A_L=20$ dB SPL. $p_{ec}[^{\omega}, A]$ from figure 6.3 is represented as a vector whose length equates to magnitude, angle subtended from in-phase axis corresponds to phase. Solid line is $p_{ec}[^{\omega}, 60]$ rescaled by $A_H-A_L$, i.e., $p_{ec}[^{\omega}, 60]$. Long dashed line is $p_{ec}[^{\omega}, 20]$. Short dashed line is the difference vector, $p_{em}[^{\omega}, A_L] = p_{ec}[^{\omega}, A_L] - p_{ec}[^{\omega}, A_H]$\(^3\). Panel a) to e), ear canal pressures at discrete frequencies. Panel f) is a vector representation over a frequency range equivalent to one cycle of microstructure.

\[
\begin{array}{cccccc}
\text{a. 1275 Hz} & \text{b. 1295 Hz} & \text{c. 1305 Hz} & \text{d. 1325 Hz} & \text{e. 1345 Hz} \\
\text{quadrature} & \text{quadrature} & \text{quadrature} & \text{quadrature} & \text{quadrature} \\
\text{in-phase} & \text{in-phase} & \text{in-phase} & \text{in-phase} & \text{in-phase} \\
\end{array}
\]

Due to the phasic interaction between $p_{em}[^{\omega}, A_L]$ and $p_{ec}[^{\omega}, A_H]$ an undulation in $|p_{ec}[^{\omega}, A_L]|$ can arise when $|p_{em}[^{\omega}, A_L]|$ is constant against frequency, corresponding in the vector representation as a disc centred on the $p_{ec}[^{\omega}, A_H]$ co-ordinate. The consequence is that the magnitude of the subtracted vector need not manifest a microstructure, whereas the real and imaginary components must possess a periodicity commensurate with $p_{ec}[^{\omega}, A_L]$.

\(^3\) Properly, $p_{em}$ is a function of both $A_L$ and $A_H$. However, it is taken that $p_{em}$ characterises the total non-linearity at $A_L$, therefore assuming $A_H$ reveals negligible non-linearity, allows for $p_{em}$ to be specified as a function of $A_L$ only.
However, experimental data (Kemp and Chum, 1980; Zwicker and Schloth, 1984) does not concur with the constant $|p_{em}[\omega, A]|$ constraint; in figure 1d and 3b from Kemp and Chum (1980), $|p_{em}[\omega, A]|$ varies by 6 dB over a couple of cycles of microstructure, a value similar to that found in data shown in figure 3 from Zwicker and Schloth (1984).

**Figure 6.11:** The subtracted sound pressure, $p_{em}[\omega, A]$, for two female subjects; SS-left column and JE-right column. Top panel shows the in-phase and quadrature components of the subtracted sound pressure in units of $\mu$Pa. Bottom panel displays the magnitude in dB SPL read from the left ordinate, and the phase in radians read from the right ordinate.
Chapter 6: Ear canal acoustic measurements

The subtraction method is very sensitive to the uncertainty on $p_{ec}'[\omega, A_H]$ and $p_{ec}[\omega, A_L]$, since the absolute error on the $p_{em}[\omega, A_L]$ is the sum of the absolute errors on $p_{ec}'[\omega, A_H]$ and $p_{ec}[\omega, A_L]$. The following section derives the subtracted sound pressure error function.

6.11 Sources of error in the subtracted sound pressure

Experimental errors relating to the CTASP measurement paradigm have two sources; one that is a function of the noise in the analysed signal, and a second which is a function of the linearity, or more specifically homogeneity, in the signal generation and measurement devices. Both errors are observable in the control data of sound pressure measurements within the B&K ear simulator, as shown in figure 6.4. The noise-dependent error is identified as a fluctuation of each datum around a mean value across frequency, i.e., deviations from a gaussian distribution, which is most evident for the low level sweep. The level-dependent error is interpreted as a vertical, gross d.c. shift in that mean value, between the two curves.

From data shown in figure 6.4, a 2.5 % shift occurs in the real part, while a 1.7 % change is observed in the argument. Taking the worst case of 2.5 %, this leads to a level dependent noise floor 32 dB less than 20 dB SPL, that is -12 dB SPL.

Vectorial representation of the sound pressures, in linear units, e.g. Pa, illustrates the construction of the error on $p_{ec}$ as a disc of uncertainty, and makes evident its role upon the uncertainty in $p_{em}$. The radius of the disc is a function of the gaussian dependent error, while the centre of the disc relative to the $p_{ec}[\omega, A_L]$ co-ordinate reflects homogeneity.

When calculating errors, it is noted that under the operation of addition (or subtraction) the total absolute error is equal to the sum of the individual absolute errors, while under multiplication (or division) the total relative error is equal to the sum of the individual errors.
relative errors\(^4\). To obtain the SFOAE, \( p_{em} \), two operations need to be performed; division, then subtraction.

Under stimulation at two different stimulus levels in linear units, hence the lower case \( a_L \) and \( a_H \), with \( a = a_H / a_L \), and the absolute error in parameter \( E \) specified as \( \Delta E \), it follows that the absolute error in \( p_{em} \) is then given by:

\[
\Delta p_{em}[\omega, a_L] = \Delta p_{ec}[\omega, a_L] + \frac{\Delta p_{ec}[\omega, a_H]}{a} + p_{ec}[\omega, a_H] \frac{\Delta a}{a}
\]

(6.10)

The first two terms on the right hand side are functions of the random fluctuations, from a gaussian distribution, of the low level and scaled-down sound pressures, and are thus a function of the noise rejection achieved by signal analysis. The third term on the right hand side determines the level dependent error that arises due to the deviation of the actual stimulus difference from the desired, specified stimulus difference, and reflects the homogeneity in the stimulus generation and measurement instrumentation.

### 6.12 Summary

The experimental results presented in this chapter support the view that microstructure, manifesting a non-linear characteristic, is a common, but not universal, phenomenon observed in physical measures of the auditory periphery. Analysis of the experimental results by the application of system functions allowed parameters, conventionally employed in the field of electrical engineering, to be estimated. Due to the phasic interactions between the stimulus and non-linear component, frequencies at which the subtracted sound pressure magnitude, \( |p_{em}| \), is greatest do not necessarily coincide with peaks in the \( |p_{ec}| \) microstructure, to such a degree that \( |p_{em}| \) is not constrained to exhibit a microstructure characteristic.

\(^4\) Consider two parameters, \( a \) and \( b \), with absolute errors \( \Delta a \) and \( \Delta b \), respectively. The absolute error in \( c = a + b \) is simply \( \Delta c = \Delta a + \Delta b \). The relative error in \( c = a \cdot b \) is given by \( c + \Delta c = (a + \Delta a)(b + \Delta b) \), ignoring the second order terms, \( \Delta c = \Delta a b + \Delta b a \), leading to \( \Delta c / c = \Delta a / a + \Delta b / b \).
Chapter 7: Derived ear canal parameters

7.1 Introduction

The undulatory nature of ear canal sound pressure against frequency, namely CTASP microstructure, evident at low, but undetectable at high stimulus levels, is generally accepted as of cochlear origin. This notion was initially proposed by Kemp (1979a) and supported by data in Wilson (1980b) and by analysis of the CTASP signal in Shera and Zweig (1993). Specifically, it is suggested that microstructure is a consequence of the presence of a hydromechanical cochlear wave travelling basally, i.e., a conventional cochlear TW propagating towards the base.

As discussed in chapter 2, one effect of a basal TW, *inter alia*, is a modification of the cochlear input impedance from that case when only apical wave propagation exists. Can the modification in the cochlear input impedance be observable as a change in the acoustic measures, or parameters, of the ear canal response?'

The following section uses network parameters such as input impedance, reflectance and power flow, that are conventionally found in the domain of electrical engineering but can equally be applied to the study of ear canal acoustics through the use of electroacoustic analogues. Subsequent sections display experimental measurements of the ear canal input impedance at two stimulus levels, and show reflectance and power flow estimates. The aim of the current chapter is to present the changes in ear canal acoustic parameters, other than sound pressure, resulting from modifications, or more specifically the nonlinearisation in cochlear input impedance.

7.2 Definition of network parameters

The acoustical character of the probe, consisting of a sound source and pressure microphone, is modelled by an equivalent Thevenin source circuit, i.e., a perfect open circuit pressure source, $p_s[\omega, A]$, and series impedance, $Z_s[\omega]$. Both Thevenin
parameters are functions of frequency, expressed here in radians per second, \( \omega \), with \( p_s \) a function of stimulus level, \( A \), in dB SPL.

The complex sound pressure sampled in the plane which is flush with the probe tip is \( p_{ec} [\omega, A] \), which through equation 2.1 allows for the calculation of ear canal input impedance at the probe tip, \( Z_{ec} [\omega, A] \). Volume velocity \( U_{ec} [\omega, A] \) has the value, \( U_{ec} = p_{ec} \mid Z_{ec} \).

### 7.3 Definition of scattering parameters for one port networks as applied to ear canal acoustics

Analysis of ear canal acoustics is facilitated by the use of scattering parameters. The two scattering parameters, \( a \) and \( b \), for a one port network given by Carlin (1956), as applied to ear canal acoustics, yield:

\[
\begin{align*}
a &= \frac{1}{2} \left( p_{ec} + Z_{ec}^a \ U_{ec} \right) \\
b &= \frac{1}{2} \left( p_{ec} - Z_{ec}^a \ U_{ec} \right)
\end{align*}
\]  

(7.1a) 

(7.1b)

where \( Z_{ec}^a \) is the reference, or wave, acoustic impedance, equal to the transmission line lossless characteristic impedance\(^3\). Analysis of ear canal acoustics specify \( Z_{ec}^a \) to be the characteristic lossless impedance of a uniform circular tube which approximates to the dimensions of the canal. \( Z_{ec}^a \) is then equal to \( \rho \ c / \pi \ a^2 \), where \( a \) is the tube radius in metres. The parameters \( a \) and \( b \) have dimensions equal to that of sound pressure. Referring to the nomenclature of Carlin (1956) the above definitions are

---

\(^{1}\) Reference to angular frequency \( \omega \), rather than linear frequency \( f \), is preferred due to the direct relationship to the vector representation of complex frequency where \( \omega \) specifies the number of radians swept per second.

\(^{2}\) Surprisingly, in contrast to the implication of the term velocity, volume velocity is a scalar quantity (Kinsler et al., 1982). The use of volume velocity rather than particle velocity preserves the analogue with mechanical speed and electrical current, thereby obeying Kirchhoff's current law.

\(^{3}\) This is usual when applying scattering parameters to distributive networks (Carlin, 1956; Kuo, 1962; Ruston and Bordogna, 1966; Davidson, 1989). However, scattering parameters can equally be applied to purely lumped immittance element networks, where \( Z_o \) is set equal to the source impedance (Carlin, 1956).
termed non-normalised, as opposed to normalised\(^4\), scattering parameters. Rearrangement of equations 7.1a and 7.1b leads to:

\[
P_{ec} = a + b \\
U_{ec} = \frac{(a - b)}{Z_{ec}^o} \tag{7.2a, 7.2b}
\]

It follows that the scattering parameters are \(a = p_{ec}^+\) and \(b = p_{ec}^-\), namely the forward and backward travelling ear canal sound pressure waves, respectively. The scattering coefficient, \(s\), is then defined as:

\[
s = \frac{b}{a} = \frac{(p_{ec} - Z_{ec}^o U_{ec})}{(p_{ec} + Z_{ec}^o U_{ec})} = \frac{Z_{ec} - Z_{ec}^o}{Z_{ec} + Z_{ec}^o} \tag{7.3}
\]

Normalisation of \(Z_{ec} [\omega, A]\) by \(Z_{ec}^o [a]\) defines the normalised ear canal input impedance, \(Z_{ec}^n [\omega, A, a]\), a dimensionless parameter, therefore:

\[
s = (Z_{ec}^n - 1)/(Z_{ec}^n + 1) \tag{7.4}
\]

Describing impedances in terms of normalised parameters allows for the simple assessment of the coupling efficiency, i.e., power transfer, between source and load.

In an ear canal acoustical network, where the parameters are defined as non-normalised, \(s\) is referred to as the ear canal pressure reflectance, \(R_{ec} [\omega, A, a]\), which is dimensionless, but a function of frequency, stimulus level and ear canal radius.

### 7.4 Power flow in a one port network as applied to ear canal acoustics

For a passive network, with \(Z_{ec}^n = r + jx\):

\[
|R_{ec}|^2 = 0 \leq \frac{(r - 1)^2 + x^2}{(r + 1)^2 + x^2} \leq 1; \quad r \geq 0 \tag{7.5}
\]

\(^4\) Normalised scattering parameters have dimensions equal to the square-root of power.
Chapter 7: Derived ear canal parameters

The dimensionless parameter $|R_{ec}|^2$, is the ear canal power reflectance, $E_{ec}^{[\omega, A, a]}$, and defines the ratio of reflected to incident power flow along the canal. If the ear canal approximates to a lossless line, where energy dissipation along the line is negligible, then $E_{ec}$ is a measure of the mismatch between the ear canal and middle ear at the TM. If $E_{ec}$ is zero then the middle ear load dissipates all the available, or incident, power. When $E_{ec}$ is equal to unity all the power delivered to the middle ear load is reflected at the TM.

The power incident, reflected and absorbed by the middle ear then is given by:-

$$\text{Power incident, } \Pi^i_{ec} = \frac{|P_{ec}^i|^2}{Z_{ec}^o} \quad (7.6a)$$

$$\text{Power reflected, } \Pi^r_{ec} = \frac{|P_{ec}^r|^2}{Z_{ec}^o} \quad (7.6b)$$

$$\text{Power absorbed, } \Pi^a_{ec} = \Pi^i_{ec} - \Pi^r_{ec} \quad (7.6c)$$

The ratio of power absorbed to power incident is then given by:-

$$\frac{\Pi^a_{ec}}{\Pi^i_{ec}} \equiv T_{ec} = 1 - E_{ec} \quad (7.7)$$

The ear canal energy transmittance, $T_{ec}^{[\omega, A, a]}$, is a dimensionless parameter and is the ratio of absorbed power to available, incident power.

7.5 Power flow constituents: an alternative scheme

If we work with complex power, $S$:-

$$S = P \pm jQ \quad (7.8)$$
where $P$ is power taken by the load, in resistive watts, and $Q$ in reactive watts, borrowing from electric circuit theory (Yorke, 1986):-

$$S_{ec} = P_{ec} U_{ec}^* = |P_{ec}|^2 / \text{Re} \{Z_{ec}\}$$

(7.9)

where $U^*$ is the complex conjugate of $U$.

With this approach, the absorbed power, $\Pi_{ec}^a[\omega, A]$, is the real part of $S_{ec}$. The incident power, $\Pi_{ec}^i[\omega, A,a]$, has the value $\Pi_{ec}^a / T_{ec}$, while the reflected power, $\Pi_{ec}^r[\omega, A,a]$, is given by $E_{ec} \Pi_{ec}^a / T_{ec}$. Note that the incident and reflected parts are functions of the radius of the ear canal.

### 7.6 Ear canal input impedance: empirical data

The ear canal input impedance, derived from experimental data using equation 2.1 and shown in figure 7.1, displays a similar microstructure characteristic to the CTASP data. Here, the low level undulating curve is superimposed upon the smooth, slowly varying high level curve, such that the periodicity of $Z_{ec}[\omega, A_L]$ is comparable to $p_{ec}[\omega, A_L]$ shown in figure 6.3. Note that the parameter $p_s[\omega, A]$ is a function of stimulus level, enabling appropriate scaling of $Z_{ec}[\omega, A]$.

---

Figure 7.1: Calculated ear canal acoustic input impedance in the plane of the probe assembly tip, $Z_{ec}[\omega, A]$, at two stimulus levels; $A = 60$ dB SPL and $A = 20$ dB SPL. Data shown for two subjects SS (left column) and JE (right column). Impedance values calculated by employing a Thevenin calibrated probe assembly and derived from ear

---

Note that, the product of sound pressure and volume velocity has dimensions equivalent to that of power (units of W), whereas, the product of sound pressure and particle velocity has dimensions equivalent to acoustic intensity (units of Wm$^{-2}$). Conventionally, power applied to an acoustic system is specified in terms of intensity, with units of Wm$^{-2}$. However, in order to allow comparison with other systems, such as various receptor systems as has been carried out by Khanna and Sherrick (1978), the more general measure of power will be used.
canal sound pressure data, $p_{ee} [\omega, A]$, figure 6.3 for SS and figure 6.5d for JE. Top panel $|Z_{ee} [\omega, A]|$ in units of dB re 1 M MKS acoustic ohm. Dotted line represents characteristic impedance of lossless transmission line of radius 3.5 mm. Middle panel shows the $Z_{ee} [\omega, A]$ argument in radians. Bottom panel displays ear canal volume velocity magnitude, $|U_{ee} [\omega, A]|$, corresponding to $A = 20$ dB SPL. Datum at every 5 Hz, with cubic spline interpolation between data points.

\[ A = 60 \text{ dB SPL} \]
\[ A = 20 \text{ dB SPL} \]
Chapter 7: Derived ear canal parameters

From figure 7.1, ear canal input impedance magnitude at 1250 Hz is equal to 23 dB re 1 M MKS acoustic ohm for subject SS, while for subject JE $|Z_{ec}[2\pi\cdot1250,60]| = 23.4$ dB, which falls close to the averaged $|Z_{ec}|$ of 25 dB for ten otologically normal adult subjects given in Keefe et al (1993).

Assessment of the acoustic drive condition, i.e. whether the power delivered to the load is developed from an approximate sound pressure source, volume velocity source or at a point between the two extremata, is made with reference to the source impedance magnitude, $|Z_s|$, relative to the ear canal input impedance magnitude, $|Z_{ec}|$; when $|Z_s| >> |Z_{ec}|$ a constant volume velocity drive condition exists; for $|Z_{ec}| >> |Z_s|$ a constant sound pressure condition exists; for $Z_s = Z_{ec}$ * the maximum power transfer criterion is met. By referring to $|Z_s|$ given in figure 5.4 and $|Z_{ec}|$ for a typical ear in figure 5.12, in the 1 to 2 kHz range, the network approximates to a constant volume velocity drive condition, under which a proportional change in $|Z_{ec}|$ leads to a equal proportional change in the ear canal sound pressure magnitude, $|P_{ec}|$.

Ear canal volume velocity magnitude, $|U_{ec}|$, given in the bottom left panel in figure 7.1, displays an undulation with frequency but with an amplitude, or proportional change, much less than $|P_{ec}|$ in figure 6.3. Peaks in $|U_{ec}|$ occur at frequencies when $|Z_{ec}|$ is at a minimum, conforming to the qualitative summary outlined above. Hence, it is shown that the drive condition approximates to a constant volume velocity source.

---

6 By analogy, in electrical terms, for constant volume velocity read constant current source, constant sound pressure read constant voltage source.
Table 7.1: Thermodynamic constants appropriate for ear canal acoustics. At a temperature of 307 °K (34 °C). From ANSI S3.25-1979. Note a -2.3% change in $\rho$ and a 1.4% change in $c$, leads to a 1% change in the characteristic impedance for plane wave propagation in air, from when temperature equalled 300 °K.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Equilibrium gas density</td>
<td>$\rho$</td>
</tr>
<tr>
<td>Speed of sound in free space</td>
<td>$c$</td>
</tr>
</tbody>
</table>

Below is the normalised ear canal input impedance, $Z_{nc}^o [\omega, A,a]$, displayed in terms of resistive and reactive parts, for $a=3.5$ mm and thermodynamic constants as given in table 7.1, corresponding to a characteristic impedance of $Z_{nc}^o [a] = 10.5$ M MKS acoustic ohm\(^7\).

Figure 7.2: Normalised ear canal resistance and reactance against frequency at two stimulus levels; $A=60$ dB SPL, $A=20$ dB SPL. Derived from $Z_{nc} [\omega, A]$ data in figure 7.1 for subject SS. Impedance normalised by characteristic impedance equal to 20.4 dB re 1 M MKS acoustic ohm, corresponding to a lossless ear canal of radius $a=3.5$ mm, estimated from the radius of ear canal for subject SS. Error bars representing 10 % uncertainty in ear canal radius for the low level case only. Datum at every 5 Hz, with cubic spline interpolation between data points.

---

\(^7\) Since the real part of impedance determines energy dissipation, it may appear counter-intuitive to find that for a lossless uniform transmission line, its characteristic impedance ($Z_o$) is real, whereas, the characteristic impedance for a lossy line consists of an additional non-zero imaginary part. However, $Z_o$ for a lossless transmission line, equivalent to the input impedance of the line when it is infinitely long, i.e., perfectly terminated, cannot be interpreted as a driving-point impedance, which is the ratio of sound pressure to volume velocity at the same point. Referring to equation 4.2, $Z_o$ is not specified at one point but is equal to the square root of the product of the series impedance per unit length and shunt admittance per unit length.
In the calculation of $Z_{ec}^n$ it is necessary to estimate the radius of the ear canal under the assumption that it acts in a similar fashion to, and can be modelled by, a straight uniform circular hard walled duct. Consequently, both the validity of the model and the estimate in ear canal radius adds uncertainty to the parameter value of $Z_{ec}^n$. The validity of the assumption made is dependent upon the stimulus frequency, since the human ear canal of an adult typically shows a curvature and variation in cross-sectional area along its length. The greatest breadth of the ear canal is at the entrance as it flares into the concha, while it tapers towards the eardrum. At high stimulus frequencies, above 8 kHz, the geometry of the ear canal results in the acoustics response deviating from that of the duct (Stinson, 1985). However, at a frequency range about 1.5 kHz such assumptions hold. Figure 7.3 shows the range in $Z_{ec}^n$ at the low stimulus level for a 10% uncertainty in ear canal radius.
Chapter 7: Derived ear canal parameters

7.7 Ear canal reflectance

By applying equation 7.3, ear canal pressure reflectance, $R_{ec}$, and subsequently energy reflectance, $E_{ec}$, were calculated and shown in figure 7.3. Ear canal energy reflectance under the high stimulus level is approximately constant with frequency, equal to 0.35, which is 70% of the figure of 0.5 given in Keefe et al (figure 4, 1993). However, the latter figure is an average and as an indication of the variance in the data, figure 17 from Keefe et al (1993) shows ear canal energy reflectance for two typical adult subjects, where $E_{ec}$ equals 0.3 and 0.56. Allen et al (1995) presents reflectance data for one adult subject, where at 1350 Hz $E_{ec} \approx 0.37$. It is concluded that the data in figure 7.3 is not atypical considering the variance in energy reflectance shown in Keefe et al (1993).

Under low level stimulation, corresponding to 20 dB SPL, microstructure is evident in all representations of reflectance. For example, $E_{ec}[\omega,20,a]$ varies between $E_{ec}[\omega,60,a] \pm 2 \text{ dB}$, a value comparable with Allen et al (1995).

Figure 7.3: Ear canal reflectance against frequency at two stimulus levels; $A = 20$ and $A = 60$ dB SPL. Derived from data in figure 7.1. Data is shown in two main blocks of four panels, top block for subject SS, and bottom block for JE. For each main block, top left panel shows $\text{Re}\{R_{ec}[\omega,A,a]\}$, bottom left panel displays $\text{Im}\{R_{ec}[\omega,A,a]\}$ and top right panel shows argument of $R_{ec}[\omega,A,a]$. Energy reflectance, $E_{ec}(\omega,A,a)$, on a logarithmic scale is shown in the bottom right panel. Both $R_{ec}[\omega,A,a]$ and $E_{ec}[\omega,A,a]$ are dimensionless parameters. Ear canal radius, $a=3.5\text{mm}$.
Chapter 7: Derived ear canal parameters

![Graphs showing derived ear canal parameters for different frequencies and sound pressure levels (SPLs). The graphs depict real and imaginary parts of the derived parameters for A=60 dB SPL and A=20 dB SPL.]

Frequency (Hz)

Re\([R_{sc}(\omega, A, a)]\)

Im\([R_{sc}(\omega, A, a)]\)

Re\([E_{sc}(\omega, A, a)]\)

Im\([E_{sc}(\omega, A, a)]\)
Chapter 7: Derived ear canal parameters

7.8 Power flow in the ear canal

This section displays data relating to the incident, reflected and absorbed constituents of ear canal power flow, calculated by applying network parameters given in equations 7.6a-c.

Figure 7.4: Ear canal power constituents, incident, reflected and absorbed, at two stimulus levels. For $A=60$ dB SPL read from left ordinate, for $A=20$ dB SPL the right ordinate. Data is shown for three subjects; SS top, JE middle and PS bottom. In order to conveniently display the data, power is given in dB re 1 fW, i.e., $10^{-15}$ W.
Chapter 7: Derived ear canal parameters

![Graphs showing derived ear canal parameters](image_url)
Chapter 7: Derived ear canal parameters

Referring to ear canal reflectance data (figure 7.3), it could be concluded that, assuming negligible power absorption along the length of the ear canal, and constant power incident, the power reflected at the TM varies by up to 3 dB across the frequency sweep. Taking the first assumption, Keefe (1984) provides a theoretical analysis of sound propagation along cylindrical tubes with reference to the transmission parameters. For a tube whose dimensions are equal to that of the standard occluded ear canal, the attenuation is found to be 0.056 dB/cm at 6 kHz. Focusing upon the second assumption, in real ear canal measurements the power incident is not constant against frequency, since the power extracted from the source is a function of $Z_{se}$, which itself exhibits a microstructure, relative to $Z_s$. A microstructure characteristic in ear canal incident power is clearly shown in figure 7.4, where at particular frequencies the ear is extracting more power from the source than at other frequencies. Therefore, due to the interaction of the probe and ear canal load, changes in the effective stimulus have to be considered when making comparisons between the dimensionless measure of energy reflectance and, for instance, absolute sound pressure, at low stimulus levels.

7.9 Summary

Experimental data of the ear canal response described in terms of impedance, reflectance and power, manifest microstructure characteristics that are observed in ear canal sound pressure, as given in chapter 6. Namely, that the amplitude of microstructure is evident at low (20 dB SPL) stimulus levels, but not detectable at higher (60 dB SPL) stimulus levels, in addition to the microstructure periodicity being commensurate with CTASP data. The suggestion that auditory microstructure originates in the presence of a basalward travelling wave is investigated in the next chapter.
Chapter 8: Cochlear parameters inferred from ear canal measurements

8.1 Introduction

The preceding two chapters present empirical data on the characterisation of the peripheral auditory system, at one measurement site within the ear canal, by four acoustic parameters; namely, sound pressure, impedance, reflectance and power flow. Due to the mechanical coupling of the ear canal to the cochlea afforded by the middle ear (ME), knowledge of the ME transformation allows some aspects of the response of the cochlea to be inferred from ear canal measurements. The ME does not respond preferentially to forward stimulation, but rather responds to both forward and retrograde stimulation equally. In other words, the ME supports bidirectional power flow and displays the same insertion loss to forward power flow as opposed to retrograde power flow.

Under the constraint of linearity, the amplitude response of the occluded ear canal, ME, the stimulus delivery and the measurement devices remains invariant against stimulus level. Cochlear non-linearity is then identifiable as a deviation from the linear case. This can be achieved experimentally through a variation in stimulus level; such a nonlinear characteristic is observable in empirical data of ear canal acoustics (Kemp, 1979a; Kemp and Chum, 1980; Wilson, 1980b; Zwicker and Schloth, 1984; Shera and Zweig, 1993; Allen et al, 1995).

The aim of the current chapter is to infer the characteristics of cochlear impedance and reflectance from the ear canal empirical data presented and from a theoretical treatment of the auditory periphery based on an elementary analogue network model.

---

1 As well as coupling in the form of mechanical energy, the tympanic cavity of the middle ear supports acoustic coupling arising from the differential pressure across the tympanic membrane. The spatial nature of the sound energy within the tympanic cavity can provide an effective acoustic stimulus to the cochlea by a pressure difference between the OW and RW (Peake et al, 1992). The nature and magnitude of acoustic coupling between the outer and inner ear is discussed later in section 8.3.3.

2 Inference, i.e., the process of reasoning through a transition of logical steps, can either take the form of deduction or induction. The former draws conclusions from principles and terminates in facts about events, and therefore proceeds from the general to the particular. Induction, in contrast, steps in the opposite direction, starting at facts about a particular event and terminates with first principles. It is argued here that deductive inference will permit discussion of the cochlear response from knowledge of
In essence, the interpretation of a level dependent component in ear canal acoustics can be made with reference to the presence or absence of a basalward cochlear, hydro-mechanical, travelling wave (TW⁺). The influence of a TW⁺, in particular its role in the modification in cochlear pressure reflectance, $R_c$, upon ear canal acoustic parameters is investigated. The following sections aim to quantitatively validate the qualitative view that $R_c$, and changes therein, are observable in ear canal acoustic measurements through the use of immittance and scattering parameters applied in a 2-port network formalism. The phenomenological formalism employed is presented in terms of acoustic parameters, but follows that developed in electrical network theory (Yorke, 1986). And it is stressed that it is only applicable if the constraints of linearity, reciprocity and one dimensional wave motion are met.

Discussion focuses upon the nature of $R_c$, against both stimulus level and frequency and its effect upon ear canal acoustics. Specifically, a cochlear pressure reflectance vector of constant magnitude but rotating uniformly with linearly increasing frequency as hypothesised by Kemp (1980) is investigated. Such a dependence upon $R_c$ is shown to be consistent with auditory microstructure observed in ear canal acoustic parameters presented in chapters 6 and 7.

8.2 Phenomenological modelling approach: 2-Port network theory

A method to model any linear system, of arbitrary topological complexity, is the $n$-port network representation (figure 8.1a). For systems in which only two pairs of independent terminals are identifiable, a 2-port network topology is appropriate (figure 8.1b). Conventionally, the input port is labelled 1-1' and the output 2-2'. 2-port networks are introduced in a lucid and didactic fashion in Yorke (1986).

---

3 The abbreviation, TW, signifies a cochlear, slow hydro-mechanical travelling wave in contrast to the fast compressional wave. TW⁺ denotes that the wave propagates apically, towards the helicotrema, whereas, TW⁻ implies that wave propagates basalward towards the stapes.
Figure 8.1: a) Representation of a generic linear n-port network, $^1\mathbf{N}_n$. b) When only two pairs of independent terminals, or ports, can be identified the formalism reduces to a 2-port network, $^1\mathbf{N}_2$. The circuit parameters $p$ and $U$ are identified as sound pressure and volume velocity, respectively, the ratio of which gives impedance, while the product is equivalent to power.

The utility of the 2-port formalism is that any linear system, active or passive, can be concisely, yet completely, described by a two-by-two matrix, known as the transmission matrix. The transmission coefficients are independent of the boundary conditions, i.e., both the input and output loads. The transmission matrix is also known by the following three terms; transformation matrix, $ABCD$ or $T$-matrix.

The transmission matrix, $^1\mathbf{T}_2$, is defined as:

$$
\begin{pmatrix}
p_1 \\
u_1
\end{pmatrix} = ^1\mathbf{T}_2
\begin{pmatrix}
p_2 \\
u_2
\end{pmatrix}
$$

(8.1a)

where:

$$
^1\mathbf{T}_2 = \begin{bmatrix} A & B \\ C & D \end{bmatrix}
$$

(8.1b)
Chapter 8: Cochlear parameters inferred from ear canal measurements

For generality, it is taken that inputs, outputs and transfer coefficients are all function of $e^{j\omega t}$, which, for the purposes of clarity, is omitted.

It follows, that four independent equations, of input and output signals, are required to solve for the four matrix elements, $A, B, C$ and $D$, known as the transfer coefficients.

Other 2-port network matrices relate the inputs and outputs in different equation configurations. There are six 2-port network matrices in all; $ABCD, abcd, Z, Y, g$ and $h$. Since all describe the same system, all are inter-related (Yorke, 1986). Anticipating the use of scattering 2-port matrices, all six matrices listed above are generically termed immittance matrices.

To fully characterise the 2-port matrix of any configuration, knowledge of the system components need not be known, only the input and output signals are measured. The 2-port network approach, therefore, lends itself conveniently to phenomenological modelling, such as 'black-box' systems, where intermediate, internal, sub-system parameters are not known. For example, in the case of acoustical systems, the 2-port transmission matrix relates input acoustic pressure ($p_1$) and volume velocity ($U_1$) in a linear combination of both output acoustic pressure ($p_2$) and volume velocity ($U_2$).

Most generally, volume velocity, as opposed to particle speed, is the acoustical parameter which is analogous to electrical current since it is conserved at nodes and therefore conforms to Kirchhoff's current law (Beranek, 1954). However, under the constraint of one dimensional wave motion, particle speed is proportional to volume velocity, the constant of proportionality given by an area factor. The ratio of sound pressure to volume-velocity gives impedance, while the product is dimensionally equivalent to power.

From equation 8.1 it follows that the acoustic input impedance at port 1-1' is interpreted as a transformation of the output load impedance, $Z_2$:

$$Z_1 = \frac{p_1}{U_1} = \frac{Z_2 \cdot A + B}{Z_2 \cdot C + D} \quad (8.2)$$
where \( Z_2 = p_2 / U_2 \).

If inputs and/or outputs are unobtainable, knowledge of the system can prove useful. Lampton (1978) lists the relationship between the topology of the system and the elements of the transmission matrix, leading to a numerical modelling approach. A reductionist paradigm can be employed enabling a multi-staged linear system to be represented by a cascade of transmission matrices, each matrix representing a simple linear stage. The transfer coefficients of the whole system can then be calculated, given simply as a matrix multiplication of all of the intermediate transmission matrices. De Boer (1980) followed this approach in order to solve for the cochlear partition velocity transfer function.

Different modes of energy transfer, electrical, mechanical or acoustical, between the input and output ports can be handled by the inclusion of an ideal transformer or gyrator. For example, modelling of the electrical (voltage, \( e \); current, \( i \)) to mechanical (force, \( F \); speed, \( u \)) transduction of a electro-dynamic moving-coil motor requires the following transducer equation pair (Beranek, 1954):

\[
F = Bli, \quad e = Blu \tag{8.3a; b}
\]

In transmission matrix form:

\[
\begin{bmatrix}
e \\
i
\end{bmatrix} =
\begin{bmatrix}
0 & Bl \\
1/Bl & 0
\end{bmatrix}
\begin{bmatrix}
F \\
u
\end{bmatrix}
\tag{8.4}
\]

where the overall electromechanical coupling constant equals the force factor, \( Bl:1 \). The transmission matrix is then of the form of a gyrator (Lampton, 1978).

Modelling linear systems with, for example, internal motion of more than one dimension, is valid so long as the input and output, sound pressures and volume velocities, give rise to one-dimensional driving point impedances; that is, impedances that can be represented by single co-ordinate vectors.
8.3 Deconstructing the peripheral auditory system

Energy flowing through the peripheral auditory system from the outer, middle to inner ears takes three forms: acoustical, mechanical, and hydro-mechanical respectively. The following deconstruction of the peripheral auditory system considers energy in the acoustical mode, with sound pressure and volume velocity as the circuit parameters analogous to the voltage-current parameter pair (Beranek, 1954). Figure 8.2 illustrates the network components and circuit variables relating to the deconstructed peripheral auditory system together with an electroacoustic probe occluding the ear canal.

Figure 8.2: Deconstructing the peripheral auditory system. Elementary analogue network incorporates a linear 2-port network, $^{em}N_e$, characterising the response of the middle ear. Electroacoustic probe occluding the ear canal modelled by its Thvenin equivalent acoustical source. The occluded ear canal is modelled as a uniform, i.e. constant cross section, transmission line. Since the cochlea displays a decreasing characteristic impedance with increasing distance from the base, it is represented as a tapering transmission line of increasing cross sectional area. Parameters with superscript $+/-$ denote forward and backward scattering parameters, respectively. Network valid only for linear, one dimensional motion.
8.3.1 Representation of the acoustic response of electroacoustic probe

From figure 8.2, the electroacoustic probe, consisting of the sound source and measurement microphone is described by its Thevenin equivalent acoustic source parameters, i.e., an ideal open circuit sound pressure source, $p_s$, coupled to a series impedance, $Z_s$. The Thevenin equivalent acoustic source parameters were obtained during the calibration procedure, as detailed in chapters 3, 4 and 5.

8.3.2 Acoustic description of the occluded ear canal

After the probe is sealed into the ear canal the remaining ear canal, of length $l_{ec}$ and area $a_{ec}$, is modelled as a uniform transmission line, of characteristic impedance $Z_{ec}^o$ and propagation wavenumber $\Gamma_{ec}$. The transmission matrix of the occluded ear canal, $\text{pN}_{mn}$, is given by (Lampton, 1978; Keefe, 1984):

$$
\text{pN}_{mn} = \frac{\cosh[\Gamma_{ec} \cdot l_{ec}]}{Z_{ec}^o \sinh[\Gamma_{ec} \cdot l_{ec}]} \begin{bmatrix} \cosh[\Gamma_{ec} \cdot l_{ec}] & Z_{ec}^o \sinh[\Gamma_{ec} \cdot l_{ec}] \\ 1/Z_{ec}^o \sinh[\Gamma_{ec} \cdot l_{ec}] & \cosh[\Gamma_{ec} \cdot l_{ec}] \end{bmatrix}
$$

(8.5)

where, for a lossless line, $Z_{ec}^o = (\rho_{ec} c_{ec})/a_{ec}$ and $\Gamma_{ec} = j\omega/c_{ec} = j2\pi/\lambda$. Table 8.1 lists five parameters for a standardised occluded ear canal. For modelling purposes, $\rho_{ec}$ and $c_{ec}$ are fixed, whereas $l_{ec}$, $a_{ec}$ and consequently $Z_{ec}^o$ are varied to match the dimensions of the subject's ear canal.
Table 8.1: Parameters of the average, standardised, occluded ear canal (ANSI S3.25-1979). Note that ambient temperature in ear canal is taken to be 34 °C.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equilibrium gas density $\rho_{ec}$</td>
<td>1.15 kg m$^{-3}$</td>
</tr>
<tr>
<td>Speed of sound in free space $c_{ec}$</td>
<td>352 m s$^{-1}$</td>
</tr>
<tr>
<td>Ear canal radius $a_{ec}$</td>
<td>3.74×10$^{-3}$ m</td>
</tr>
<tr>
<td>Ear canal characteristic impedance $Z'_{ec}$</td>
<td>9.2 M MKS acoustic ohm</td>
</tr>
<tr>
<td>Length of occluded ear canal $l_{ec}$</td>
<td>1.26×10$^{-2}$ m</td>
</tr>
<tr>
<td></td>
<td>19.2 dB re M MKS acoustic ohm</td>
</tr>
</tbody>
</table>

At the probe tip the occluded ear canal sound pressure and volume velocity are given by $p_{ec}$ and $U_{ec}$, respectively, the ratio of which results in the ear canal input impedance, $Z_{ec}$.

8.3.3 Modelling the mechanics of the middle ear complex

Referring to figure 8.2, the occluded ear canal is terminated by a 2-port network, $^{\text{in}}N_e$, which describes the response of eardrum together with the ossicular chain and enclosed tympanic cavity. $^{\text{in}}N_e$ is loaded by an impedance, $Z_e$, representing the cochlear input impedance. Thus, the ear canal acoustic environment is made explicitly a function of the middle ear loaded by the cochlea.

The action of the ME is commonly referred to as an impedance transformation (von Békésy, 1960). An ideal transformer has zero capacity for energy dissipation, i.e., lossless, and is generically known as a positive immittance converter, a term which reveals the transformation: with a turns ratio of $n:1$, the load impedance on port 2-2’ is modified by $n^2$ at the 1-1’ port (Lampton, 1978). The transmission matrix representing an ideal transformer is given in equation 8.4.
A transformation from the characteristic impedance of the cochlear fluid to that of air, such that the energy loss of the compressional wave in both fluids is minimised, is incorrect for the following reasons. The effective stimulus is not the launching of a compressional wave within the cochlear fluid, which equates to the fast wave as discussed in section 2.7, but rather the initiating of a slow, hydro-mechanical, travelling wave (Zwislocki, 1950). The effective cochlear driving point impedance is therefore a function of the fluid flow between the OW to RW and its interaction with the mobile cochlear partition. Hence, the net pressure across the cochlear partition, i.e., the transpartition sound pressure, at the base initiates the TW. Hence, to ensure efficient energy transfer, the function of the middle ear is to couple, or transform (equation 8.2), the cochlear input impedance to that of the characteristic impedance of the ear canal.

If the middle ear acted as an optimal coupling device, achieved through the use of an ideal transformer of appropriate turns ratio, the input impedance at the TM would equal the ear canal characteristic impedance, $Z_{ce}^p$. Since under such a condition no reflection would occur at the TM, the ear canal energy reflectance would be zero, a finding which is not observed in experimental data at high stimulus levels (figure 7.4; Keefe et al, 1993). Although the use of a transformer which optimally couples the cochlea to the ear canal is inappropriate, modelling the ME as a transformer with a turns ratio equal to 1/30, which results in realistic values for ear canal energy reflectance, will be used because of its simplicity.

Zwislocki (1962) proposed a quantitative model of the middle ear, developed from measurements on pathological ears and anatomical data, and represented as a network of electrical elements based upon functional anatomy. The measure employed as 'the goodness of fit' between model predictions and experimental measurements was the acoustic input impedance at the TM. The link between middle ear mechanisms and circuit topology is shown in figure 2 (Zwislocki, 1962). Due to the position of the block representing the effect of middle ear cavities, which precedes the blocks associated with the ossicles, it is made explicit that the cochlea can only be stimulated through the transmission path effected by motion of the ossicular chain.
Chapter 8: Cochlear parameters inferred from ear canal measurements

Utilising the formalism given in Lampton (1978) which provides a link between network topology and the transfer matrix coefficients, the Zwislocki middle ear, less the cochlear complex branch is represented as a single 2-port network (figure 8.3). Parameter values are those given in Zwislocki (1962).

**Figure 8.3:** Network topology of Zwislocki analogue network of middle ear less the stapes/cochlear complex branch, relating each block on the basis of functional anatomy. $Z_1$, impedance representing middle ear cavities, $Z_2$ eardrum, $Z_3$ eardrum-malleus-incus complex and $Z_4$ incudo-stapedial joint. $Y_n$ is the driving point admittance given by $1/Z_n$.

![Network topology diagram](image)

$$
\begin{align*}
T^m_c &= \begin{bmatrix}
1 & Z_1 \\
0 & 1
\end{bmatrix}
\begin{bmatrix}
1 & 0 \\
Y_2 & 1
\end{bmatrix}
\begin{bmatrix}
1 & Z_3 \\
0 & 1
\end{bmatrix}
\begin{bmatrix}
1 & 0 \\
Y_4 & 1
\end{bmatrix}
\end{align*}
\tag{8.6}
$$

A major drawback of the models derived from Zwislocki's ME network topology is the failure to model ears in which a dislocation or complete interruption of the ossicular chain occurs. This pathway is termed ossicular coupling, while the mechanism for the initiation of a cochlear hydro-mechanical wave via an air-borne sound pressure difference across the OW and RW is identified as acoustic coupling (Peake *et al*, 1992). The latter mechanism explains the less than profound hearing loss experienced by subjects with dislocated or interrupted ossicular chain pathologies, which is further ameliorated by creating a greater pressure difference between the OW and RW, a surgical procedure known as type IV Tympanoplasty (Peake *et al*, 1997). However, studies indicate that the acoustic coupling provides little or no effect, of the order of 40 dB below ossicular coupling in terms of pressure, for otologically normal subjects (Voss *et al*, 1996).
Chapter 8: Cochlear parameters inferred from ear canal measurements

The contraction of the stapedius muscle, which is attached to the stapes, occurs during high intensity sound input; this phenomenon is known as the acoustic reflex. A study of the acoustic reflex using an extension of Zwislocki's model is developed in Lutman and Martin (1979), whereby the change in stiffness due to stapedial contraction is modelled by an additional element, a variable capacitor, $C_v$; under no reflex $C_v$ is negligibly small, while the reflex is present $C_v$ is increased.

Due to the simplicity of the ideal transformer and the common use of the Zwislocki analogue model, both networks will be used to model $N_e$.

8.3.4 Range of validity of the proposed model

The range, in frequency, of validity of the model presented is limited to when one dimensional motion ceases along the occluded ear canal, at the TM, the stapes footplate which is directly attached to the OW and within the cochlear fluid at the base. As stated in section 8.2, a 2-port description of ossicular motion in three dimensions is valid so long as the impedances at the ports can be described by one dimensional, single-vector driving-point impedances.

Since the equations governing air-borne acoustic wave motion within the audio frequency range and physiological intensities are linear (Kinsler et al., 1982), together with plane wave propagation dominating motion within the ear canal (Stinson, 1985), 2-port formalism can be applied to characterise the ear canal.

Non-planar modes do exist at the concha and TM (Shaw, 1974; Stinson and Shaw, 1982). However, sealing the ear canal with an electroacoustic probe results in non-planar modes when the occluded ear canal is driven, but within the frequency range of interest, are negligible with respect to the propagating mode, as discussed in section 4.6 and Brass and Locke (1997). Rabbitt and Friedrich (1991) provide analytical and numerical analysis of the distribution of sound pressure in the human ear canal and state that for normal sized adult ear canals the planar mode dominates.
Attached directly to the TM is the manubrium, the process of the malleus. As a result of the manubrium pulling medially, the TM is conical. Therefore, motion of the TM cannot be considered piston-like, a condition under which particle displacement is uniform across its area. Von Békésy (1960), using a capacitive probe technique, demonstrated that motion of the TM was more akin to a stiff plate hinged on its side, leading to the concept of an effective area. Using time-averaged holography and fresh human cadaver specimens, Tonndorf and Khanna (1972) reported that the centre of effort, i.e., the point of greatest excursion, did not coincide with the tip of the manubrium, but varied as a function of frequency. However, over a frequency range covering a small number of microstructure cycles, the centre of effort, and therefore the ME transfer function, remains constant.

With respect to stimulus intensity, the validity of the model is governed by linearity. As stated in section 2.10, nonlinearity in the response of the middle ear exists at high stimulus intensities where the motion of the ossicles, in particular the stapes, is asymmetric (Møller, 1974). At lower stimulus intensities, and below the threshold of the middle ear reflex, direct observations indicate that the middle ear response is linear, for example Rubinstein et al (1966).

Reciprocity is guaranteed in all linear passive systems (Rayleigh, 1896; Yorke, 1986). Thus, the middle ear manifests a reciprocal characteristic, which in terms of 2-port network theory states that the transfer impedance, i.e., the ratio of the output sound pressure to input volume velocity remains unchanged if the output and input ports are interchanged.
8.4 Cochlear input impedance inferred from ear canal input impedance

This portion of the chapter is concerned with the relationship between \( Z_{sc} \), derived from ear canal sound pressure measurements, as given in chapter 7, and \( Z_c \). Transfer coefficients relating to the intermediate network, \( pN_c \), a result of cascading \( pN_{nm} \) and \( mN_c \), are not known. However, using the assumptions outlined in section 8.3, where \( pN_{nm} \) approximates to a uniform transmission line and \( mN_c \) either an ideal transformer or Zwislocki middle ear, enables \( Z_c \) to be expressed as a function of \( Z_{sc} \):

\[
Z_c = \frac{aZ_{sc} + b}{cZ_{sc} + d} \quad (8.7)
\]

where, \( abcd \) are the inverse transfer coefficients (Yorke, 1986) of \( pN_c \). Note that, in the following treatment the frequency dependent term, \( \omega \), is omitted for clarity, and is therefore valid only for harmonic stimulation at one particular angular frequency.

With the choice of middle ear network made, the occluded ear canal parameters now need to be estimated. For subject SS, \( a_{sc} \) was set to 3.5 mm, equal to the value used to derive \( Z_{sc}^2 \), \( R_{sc} \) and \( \Pi_{sc} \) in chapter 7. For reasons developed in section 8.5.2 the length of the occluded ear canal was selected to minimise \( \text{Im}\{Z_c(\omega, A^H)\} = 0 \). Under this condition, \( l_{sc} = 1.9 \) cm for the transformer and \( l_{sc} = 1.6 \) cm for Zwislocki analogue model.
Figure 8.4: Cochlear input impedance inferred from ear canal input impedance at two stimulus levels; A=60 dB SPL and A=20 dB SPL. Ear canal input impedance (figure 7.1) derived from ear canal sound pressure (figure 6.3) for subject SS. Calculated using analogue network of peripheral auditory system (figure 8.2) for two middle ear networks; ideal transformer (left column), and Zwislocki middle ear analogue network (right column). Top panel acoustic impedance magnitude in dB re 1 M MKS acoustic ohms, and bottom panel impedance argument in radians. Data points at 5 Hz intervals with cubic spline interpolation.

By comparing ear canal input impedance (figure 7.1) derived from sound pressure measurements (figure 6.3) with inferred cochlear input impedance (figure 8.4), four observations can be made.
Firstly, in common with acoustic measures of the auditory peripheral system presented so far, the high level ($A=60$ dB SPL) cochlear input impedance remains almost flat against frequency, whereas, the low level ($A=20$ dB SPL) displays a microstructure periodicity commensurate with ear canal parameters.

Secondly, on the average, $|Z_c|$ is $67$ dB greater than $|Z_{ec}|$, a factor equivalent to a turns ratio of $n = 1/47$; almost half of the value attributed to the impedance conversion effected by the transformer, where $n = 1/30$. The difference is accounted for by the effective transforming action of the occluded ear canal. Consider the simplified case of one dimensional wave motion within a hard walled rigidly terminated transmission line. Here, the terminating impedance is infinitely high, yet at a distance equal to an odd order integer multiple of the quarter wavelength distance, the input impedance, assuming a lossless line, tends to zero.

Thirdly, the most striking feature between $Z_c$ and $Z_{ec}$ is the magnitude, or ratio of maxima to minima, of the microstructure; $|Z_{ec}|_{\text{max/min}} \approx 3$ dB, while $|Z_c|_{\text{max/min}} \approx 8$ dB. The action of the middle ear may be thought to contribute to such a reduction effect. However, if the response of the ME is primarily an impedance transformation, then proportional changes in $Z_c$ equal proportional changes in the input impedance at the TM. It is the role of the occluded ear canal that determines the reduction effect.

Fourthly, due to the choice of value of $l_{ec}$, the argument of $Z_c$ is, on the average, close to zero, indicating an almost purely resistive driving point impedance. At the lower stimulus level, the argument undulates about zero. However, the degree of microstructure does not result in a negative resistance part, since $\pi/2 < \angle R_c < \pi/2$.

**8.5 Physical basis of the origin of microstructure**

To facilitate discussion on the nature of the cochlear input impedance reference is made to a macro-mechanical description of the propagation of the TW.
Since the wavelength of the TW is small relative to the radius of curvature of the cochlear spiral, uncoiling the cochlea does not influence the motion of the TW (Zweig, 1991) and provides a simplified geometry allowing easy identification of the three fluid channels, or scalae; the scala vestibule (SV), scala tympani (ST) and scala media (SM). Reissener’s membrane, which separates the SV and SM, is highly compliant, enabling the fluid in both scalae to be considered to act as one fluid channel (Lighthill, 1991), termed the upper channel. The lower channel is set equal to the ST. The upper and lower fluid channels, separated by the cochlear partition, then approximate the gross macromechanical features of the mammalian cochlea.

### 8.5.1 Theoretical treatment of cochlear input impedance

Zwislocki (1950) and Peterson and Bogert (1950), in response to von Békésy’s (1960) discovery that the cochlear supports travelling waves as opposed to an array of resonators as proposed by Helmholtz (1863), developed a transmission line formalism commonly found in the domain of electrical engineering (Davidson, 1992). Application of the transmission line formalism is only valid under the constraint of one dimensional fluid flow, in which the particle velocity of the fluid is uniform across the cross section of the channels. The one dimensional fluid flow constraint implies that the TW wavelength is greater than the height and width dimensions of the channels; referred to as the long-wave approximation (de Boer, 1980). In the basal region, the long wave approximation is valid, whereas close to the characteristic place the approximation is violated since the TW slows down and consequently the wavelength tends to zero (Zweig et al, 1976).

The transmission line formalism defines two parameters to characterise the mechanical response of the cochlear duct at a distance from the base; a series impedance per unit length, \( Z_p \), and shunt admittance per unit length, \( Y_{cp} \). Inertial properties of the fluid

---

4 Strictly speaking, the Reticular lamina provides the demarcation between the SM and ST, acting as ionic barrier. However, the increase in fluid contained in this additional volume is small, especially at the base where the cross sectional area of the SV plus ST dominate the total cross sectional area of the cochlear duct.
are carried in $Z_F$, while $Y_{CP}$ is specified by the cochlear partition driving point admittance.

Under harmonic stapedial motion, $\omega$, and an axial distance $x$ from the base, the ratio of the transpartition sound pressure to volume velocity in the cochlear fluid specifies the input impedance. Therefore at $x = 0$, the cochlear input impedance is defined as:-

$$Z_c[\omega, A] = \left. \frac{P_c[\omega, A]}{U_c[\omega, A]} \right|_{x=0}$$

8.5.2 High level linear regime

During stimulation at moderately high levels, ~60 dB SPL, the influence of TW wave is negligible due to the lack of microstructure in ear canal parameters; it follows that the TW wave dominates wave propagation along the cochlea (Chapters 6 and 7; Kemp, 1979ab; Kemp and Chum, 1980; Zwicker and Schloth, 1984; Shera and Zweig, 1993). From now on, $A_H$ specifies a stimulus level within the high level linear regime, bounded on the lower limit by cochlear nonlinearity while the threshold of the acoustic reflex determines the upper limit.
Chapter 8: Cochlear parameters inferred from ear canal measurements

Assuming that the cochlea manifests scaling symmetry, direct measurements of BM velocity indicate that under harmonic stimulation of mid frequencies and stimulus level $A_H$, the phase response changes relatively slowly in the first turn of the cochlea (Rhode, 1971; Robles et al, 1986; Zweig, 1991; Brass and Kemp, 1993). This implies that only a small decrease in the wavelength occurs as the TW+ progresses from the base to the second turn, as well as a smoothly changing characteristic impedance.

Under the long wave approximation together with a slowly changing mechanical property in the basal turn and absence of the TW−, the high level cochlear input impedance defined in equation 8.8 approximates to the characteristic impedance of the cochlear duct (Peterson and Bogert, 1950; Zweig et al, 1976; Shera and Zweig, 1991a):

$$Z_c[\omega, A_H] \approx Z_c^\circ[\omega, x]_{x=0} \quad (8.9)$$

From transmission line theory, the characteristic impedance at $x$ is given by (Davidson, 1992):

$$Z_c^\circ[\omega, x] = \sqrt{Z_F[\omega, x] / Y_{CP}[\omega, x]} \quad (8.10)$$

Since the cochlear fluid is assumed to be inviscid, the series impedance per unit length is a function of the inertance reactance only:

$$Z_F[\omega, x] = j\omega M_F[x] = j\omega \rho / S[x] \quad (8.11)$$

where $\rho$ is fluid density and $S[x]$ the cross sectional area at $x$:

$$S[x] = \frac{S_v[x]S_t[x]}{S_v[x] + S_t[x]} \quad (8.12)$$
Chapter 8: Cochlear parameters inferred from ear canal measurements

where $S_v[x]$ and $S_i[x]$ are the cross-sectional areas of the SV and ST, respectively. With the cochlear partition impedance described by a harmonic resonator (de Boer, 1980):

$$Z_{cp}[\omega, x] = R_{cp}[x] + j\omega M_{cp}[x] + 1/j\omega C_{cp}[x]$$ \hspace{1cm} (8.13)

the shunt admittance per unit length is given by an orthogonal sum of conductance, $G_{cp}$, and susceptance, $B_{cp}$, per unit length:

$$Y_{cp}[\omega, x] = G_{cp}[\omega, x] + jB_{cp}[\omega, x]$$ \hspace{1cm} (8.14)

In the region of the base, the stiffness of the cochlear partition impedance dominates $Z_{cp}$ (von Békésy, 1960; Gummer et al, 1981):

$$Y_{cp}[\omega, x] = j\omega C_{cp}[x]_{basal\ turn}$$ \hspace{1cm} (8.15)

Substituting equation 8.11 and 8.15 into 8.10 leads to:

$$Z^*_{cp}[\omega, x] \approx \frac{M_{p}[x]}{C_{cp}[x]_{basal\ turn}}$$ \hspace{1cm} (8.16)

Compliance of the cochlear partition, which is primarily governed by the width of the basilar membrane, decreases approximately exponentially from the base to apex by nearly four orders of magnitude (von Békésy, 1960; Gummer et al, 1981), therefore:

$$C_{cp}[x] = C_o b_o \exp[\alpha b_x]$$ \hspace{1cm} (8.17)

where the width of the cochlear partition at the base is $b_o$, and $C_o$ the compliance constant.
Similarly, allowing the series inertance to vary exponentially with distance:-

\[ M_F(x) = \frac{\rho}{S_o \exp[\alpha_s x]} \]  

(8.18)

where \( S_o \) is the scala area constant.

In order to satisfy the observation that the TW manifests a small decrease in wavelength as it propagates through the basal turn, \( \alpha_s = -\alpha_b \). In other words \( Z_F \) and \( Y_{IP} \) co-vary to maintain proportionality; as \( x \) increases from base to the second turn, basilar membrane width is inversely proportional to scala area. From equations 8.17 and 8.18 it follows:

\[ Z_c^e(\omega, x) \approx \sqrt{\frac{\rho}{S_o b_o C_o}} \text{basal turn} \]  

(8.19)

With parameter values, \( \rho = 1000 \text{ kg m}^{-3} \), \( S_o = 10^{-6} \text{ m}^2 \), \( b_o = 10^{-4} \text{ m} \), \( C_o = 10^{-10} \text{ kg}^{-1} \text{ m}^2 \text{ s}^2 \), the cochlear input impedance is approximately \( 3.2 \times 10^{11} \text{ MKS acoustic ohms} \) (110 dB re M MKS acoustic ohm).

Equation 8.19, which is real and independent of frequency, is equivalent to the asymptotic expression of the exact analytical solution for the cochlear input impedance (de Boer, 1980). To reiterate, the approximate expression given in 8.19 is valid for mid frequencies, during stimulation levels corresponding to \( A_H \).

For reasons of convenience the cross-sectional area of the cochlear duct is often set independent of axial distance, \( x \), i.e., \( \alpha_s = 0 \), leading to a ‘box-model’ approximation (de Boer, 1980). Under the box-model approximation, variations in the mechanical properties of the cochlear duct per unit length are carried in the cochlear partition impedance, \( Z_{CP} \), only, since the fluid channels are isotropic. Shera and Zweig (1991a) argue that the box-model violates the slowly changing wavelength constraint upon the
TW propagation; for the box-model design, since $\alpha_x \neq \alpha_y$, equation 8.19 then becomes a function of $x$.

Zwislocki (1962) noted that the cochlear input impedance, $Z_c$, is dominated by a resistive part, a condition which is favourable, since ideally the function of the cochlear is to absorb power, concurring with experimental observations in the cat (Nedzelnitsky, 1980; Lynch et al, 1982). The study by Zwislocki (1962) was prior to any discussion of the existence of a basalward TW and its role in modifying ear canal input impedance, and is therefore consistent for moderate to high stimulus levels. Taking into account the ME transformation, Zwislocki (1962) used a figure of $2.9 \times 10^{10}$ MKS acoustic ohms (89 dB re 1 M MKS acoustic ohms) to characterise the resistive part of $Z_c$. Von Békésy (1960) identified a small mass component in $Z_c$ which Zwislocki (1962) attributed to the bulging volume known as the vestibule located immediately interior to the OW and noted that its effect in the analogue model proved negligible.

Data on the cochlear input impedance in the cat indicates that between 0.5 and 5 kHz, $|Z_c| \approx 10^{11}$ MKS acoustic ohms (100 dB re 1 M MKS acoustic ohms) with an argument non-systematic variation about zero of roughly 20 degrees (Lynch et al, 1982).

### 8.5.3 Low level linear regime: Role of cochlear pressure reflectance

At the base, the cochlear pressure reflectance, $R_c$, is defined as a function of frequency and stimulus level, in an analogous fashion to ear canal pressure reflectance:-

$$R_c[\omega, A] = \frac{p_c^-[\omega, A]}{p_c^+[\omega, A]_{x=0}} \quad (8.20)$$

where $p_c^+$ is the forward, apically travelling transpartition pressure wave and $p_c^-$ the backward, basally travelling transpartition pressure wave at $x=0$.\
From equation 7.3, and expressing cochlear input impedance in terms of cochlear reflectance, reveals an additional term for the case when $R_c \neq 0$, i.e., equation 8.9:-

$$Z_e[\omega, A_L] = Z_e^0[\omega, 0] \frac{1 + R_e[\omega, A_L]}{1 - R_e[\omega, A_L]} \quad (8.21)$$

From equation 8.21, the basal cochlear characteristic impedance is perturbed by $R_c$, where $A_L$ denotes a stimulus level within the low level linear regime.

### 8.5.4 Nature of cochlear pressure reflectance

However, what form does $R_e[\omega, A]$, a complex quantity, take with increasing frequency? Kemp (1980) postulated that, for increasing frequency, a cochlear pressure reflectance vector of constant magnitude, rotating uniformly with linearly increasing frequency could explain the change in cochlear input impedance. This conjecture is observed in the acoustics of the simplified case of uniform transmission lines.

For a rigidly terminated hard walled line, under plane wave stimulation, it is observed that the input impedance magnitude exhibits a series of maxima and minima with frequency, the spacing of which is a function of line length. However, the energy reflectance, and therefore the pressure reflectance magnitude is monotonic (Keefe et al, 1993) being a function of the real part of the terminating impedance, because the line is lossless. The origin of the undulating input impedance arises from phasic interaction between backward and forward travelling pressure waves, resulting in a pressure reflectance vector that rotates clockwise, with constant magnitude against increasing frequency.

Shera and Zweig (1993), in a study of the travelling wave ratio, a term synonymous with $R_e$, concluded that the cochlear pressure reflectance characteristic originates predominately through such a non-periodic phase function of frequency, thus concurring with, and providing evidence for, Kemp’s (1980) hypothesis.
Figure 8.5: Model cochlear and ear canal input impedance as a function of cochlear pressure reflectance. Top left panel: Cochlear input impedance, $Z_c$, calculated using equation 8.21. Top right panel: Cochlear input impedance microstructure magnitude, ratio of maxima to minima, as a function of $|R_c|$, where $|Z_{c_{\text{max/min}}}| = ((1+|R_c|) / (1-|R_c|))^2$. Ear canal input impedance calculated using elementary analogue network of peripheral auditory system (figure 8.2) with standard ear canal parameters (table 8.1) with a) ideal transformer (middle left panel) and b) Zwislocki middle ear analogue network (middle right panel). Cochlear pressure reflectance against frequency described by vector of constant magnitude, rotating with a spectral periodicity of $1/15\;f$. Curves plotted against six cochlear pressure reflectance magnitude values; 0 to 1.25 in 0.25 steps, annotated a-f. Bottom panel displays the ratio of $|Z_{ce}|$ to $|Z_{nm}|$ microstructure magnitude in dB.
To illustrate the reduction in the microstructure magnitude observed by comparing figures 7.1 and 8.4, figure 8.5 displays model values of $Z_c$ and $Z_{xc}$ as a function of cochlear pressure reflectance. From equation 8.21, $R_c$ effectively perturbs $Z_c[i\omega, A_H]$, a consequence of which is an infinite cochlear input impedance microstructure magnitude for when $R_c = 1 + i0$, as seen in the top right panel. Determination of
cochlear pressure reflectance magnitude may seem problematic since $|R_c|$ is two-valued for a given $|Z_{c\text{max/min}}|$. However, for $|R_c|<1$, the real part of cochlear input impedance is positive, while for $|R_c|>1$, $\text{Re}(Z_c)<0$, enabling $|R_c|$ to be determined.

Using equation 8.2, the ear canal input impedance microstructure magnitude is then a function of $|R_c|$; e.g., $|R_c(\omega, A_H)|=0$. For $|R_c|=0.25$, $|Z_{c\text{max/min}}| \approx 4$ dB, while $|Z_{c\text{max/min}}| \approx 8$ dB. This reduction in microstructure magnitude indicates that the location at the probe tip is not directly coupled to the TM but, due to the volume of the occluded ear canal, is decoupled, or shunted by a capacitive element.

Noting an 8 dB undulation in the inferred $|Z_c|$ (figure 8.4), the model data of $|Z_c|$ as a function of $R_c$ suggests that the magnitude of the cochlear pressure reflectance is approximately 0.25 (figure 8.5), implying that the backward travelling energy is 1/15th of the forward travelling energy.

Figure 8.5 evinces that, through the use of a elementary linear 2-port analogue network of the peripheral auditory system, $R_c$, described by a vector of constant magnitude and uniform angular velocity, can model the gross features of the empirical ear canal input impedance.

### 8.6 Two-port network model incorporating scattering parameters

Section 8.3 applied a 2-port network formalism to model components of the auditory periphery, with respect to sound pressures and volume velocities at the two ports. The resulting matrices were of immittance type (Lampton, 1978). Six scattering matrices, that relate the forward and reverse sound pressures at both ports in a linear combination, can be determined if the characteristic impedances on both ports is known (Carlin, 1956; Chen, 1975). Converting the 2-port network $^6\text{N}_c$ to a scattering formalism lends itself to the study of wave reflection and transmission, allowing the ear canal reflectance to be specified as a function of cochlear pressure reflectance (Carlin, 1956; Shera and Zweig, 1992).
With reference to figure 8.2, one of the six possible scattering matrices is defined as:-

\[
\begin{pmatrix}
  p_{ec}^+ \\
  p_{ec}^-
\end{pmatrix} = pS_e \begin{pmatrix}
  p_{ec}^+ \\
  p_{ec}^-
\end{pmatrix}
\]  

(8.22a)

where:-

\[
pS_e = \begin{bmatrix}
  r^+ & t^- \\
  t^+ & r^-
\end{bmatrix}
\]  

(8.22b)

with the scattering coefficients defined as:-

\[
r^+ \equiv \frac{p_{ec}^-}{p_{ec}^+} \Bigg|_{p_e' = 0}; \quad t^+ \equiv \frac{p_{ec}^+}{p_{ec}^-} \Bigg|_{p_e' = 0}; \quad r^- \equiv \frac{p_{ec}^+}{p_{ec}^-} \Bigg|_{p_e' = 0}; \quad t^- \equiv \frac{p_{ec}^-}{p_{ec}^-} \Bigg|_{p_e' = 0}
\]  

(8.22c-f)

The scattering coefficients are complex quantities, each a function of angular frequency, \( \omega \). Since \( pS_e \) characterises the network \( pN_e \), which is assumed to be linear, equations 8.22 are independent of stimulus level, \( A \).

When the port 2-2’ is perfectly matched, i.e. \( p_{ec}^- = 0 \), then \( r^+ \) and \( t^+ \) are the forward reflection and transmission coefficients, respectively. Similarly, for the special case when the load on port 1-1’ is perfectly matched, \( p_{ec}^+ = 0 \), then \( r^- \) is the reverse reflection coefficient, and \( t^- \) the reverse transmission coefficient.

For the general case of non-optimally matched loads, the reflection and transmission coefficients are not given by \( r \) and \( t \), but rather are given by the effective scattering coefficients.

The boundary condition at the 2-2’ port is given by the reflection coefficient:-

\[
p_2(\omega, A) = \frac{p_{ec}^-(\omega, A)}{p_{ec}^+(\omega, A)} = R_c(\omega, A)
\]  

(8.23)
Chapter 8: Cochlear parameters inferred from ear canal measurements

equivalent to the cochlear pressure reflectance. Similarly, the reflection coefficient due
to the mismatch of the load terminating the 1-1' port, defines the second boundary
c condition:-

\[ \rho_1(\omega, A) = \frac{P_{\text{in}}(\omega, A)}{P_{\text{out}}(\omega, A)} \]  
(8.24)

The effective forward reflection coefficient is:-

\[ r_{\text{eff}}^+ (\omega, A) = r^+ + \frac{t^+ t^- \rho_2(\omega, A)}{1 - r^- \rho_2(\omega, A)} \]  
(8.25a)

Similarly, the effective reverse reflection coefficient is given by:-

\[ r_{\text{eff}}^- (\omega, A) = r^- + \frac{t^+ t^- \rho_1(\omega, A)}{1 - r^+ \rho_1(\omega, A)} \]  
(8.25b)

While the effective forward transmission coefficient is defined as:-

\[ t_{\text{eff}}^+ (\omega, A) = \frac{t^+}{1 - r^- \rho_2(\omega, A)} \]  
(8.25c)

And the effective reverse transmission coefficient is:-

\[ t_{\text{eff}}^- (\omega, A) = \frac{t^-}{1 - r^+ \rho_1(\omega, A)} \]  
(8.25d)

From equation 8.25a, for stimulus levels commensurate with the high level linear
regime, \( R_c \) tends to zero, resulting in \( R_{\text{ac}} \) equal to the forward pressure reflectance
scattering coefficient, \( r^+ \) :-
Since \( r^+ \) is a function of the intermediate network, i.e., the occluded ear canal cascaded with the middle ear, \( R_{ee}[\omega, A] \) reveals a smooth, slowly changing characteristic lacking the rapidly varying microstructure manifest at low stimulus levels, evident in empirical data given in chapter 7, Keefe et al (1993), Shera and Zweig (1993) and Voss and Allen (1994).

As \( A \rightarrow A_L \), \( R_{ee} \) is the sum of \( r^+ \) and \( R_e \) as seen on the 1-1' port. The latter part governs, and contributes to, the degree of microstructure observed in \( R_{ee} \). Rearrangement of equation 8.25b expresses \( R_e \) as a function of \( R_{ee} \):

\[
R_c[\omega, A] = \frac{R_{ee}[\omega, A] - r^+}{r^- R_{ee}[\omega, A] + t^+ t^- - r^+ r^-} \quad (8.29)
\]

8.7 Inferred cochlear pressure reflectance

Applying the 2-port scattering formalism outlined in section 8.6 to the elementary analogue network model of the peripheral auditory system (figure 8.2), enables the cochlear pressure reflectance to be inferred from the derived ear canal pressure reflectance through equation 8.29. Parameter values were determined in previous sections; \( a_{ec} \) was estimated in section 7.3 in deriving \( Z_{ec}^n \); value of \( l_{ec} \) selected in section 8.4 to infer \( Z_e \).
Chapter 8: Cochlear parameters inferred from ear canal measurements

Figure 8.6: Inferred cochlear pressure reflectance, $R_c$, from derived ear canal pressure reflectance, $R_{ec}$. Data shown for two subjects; SS (left column) and JE (right column). For each subject, data is shown in cochlear energy reflectance, i.e. $|R_c|^2$ (top panel), argument (middle panel) and imaginary against real part (bottom panel). Argument data annotated with slope of the best-fit line in ms, i.e., $d\angle R_c/d\omega$ and at the centre frequency the number of waves of delay. $R_c$ inferred from derived $R_{ec}$ (figure 7.3) using equation 8.29 with analogue model of the peripheral auditory system (figure 8.2) incorporating Zwislocki model of ME. Curves correspond to a stimulation level $A = 20\,\text{dB SPL}$. Data points every 5 Hz (SS) and 10 Hz (JE).
From figure 8.6, inferred $\angle R_c$ exhibits an approximate periodicity, due to the continuous sweeping of the vector clockwise, at an almost uniform rate against frequency. The behaviour of the cochlear pressure reflectance argument with increasing frequency concurs with the character of $\angle R_c$, as proposed by Kemp (1980) and observed by Shera and Zweig (1993).
The almost linear behaviour of $\angle R_e$ against increasing frequency is characteristic of a system with an inherent propagation delay, suggesting that the slope of $\angle R_e$ carries information on the propagation time between the base and the apical reflector. The propagation time from the base to the effective reflector and back again is given by the negative of the phase slope, which for a given frequency can be expressed in terms of number of waves.

In contrast to a $|R_{ce}|$ characteristic invariant with frequency, the inferred cochlear pressure reflectance magnitude for JE displays a variation against frequency, or microstructure. Comparing $|R_{ce}|$ (figure 7.3) with $|R_e|$ (figure 8.6), the microstructure magnitude is diminished in the latter. The microstructure observed in $|R_e|$ may be a consequence of stimulating the ear at a point on the input output curve in which a small nonlinear component is present, rather than within a purely linear regime. At the stimulus level $A = 20$ dB SPL the ear canal response is predominately linear, but at stimulus levels above manifests a saturating characteristic (figure 6.7).

Consider low level acoustic stimulation of the occluded ear canal swept across frequency. Once steady state is attained, power absorbed at the TM, and consequently by the cochlea, varies with frequency (Figure 7.4). Within a purely linear regime, $|R_e|$ remains invariant to changes in power absorbed. At a stimulus level corresponding to the transition between the low level linear regime and mid level nonlinear regime, the nonlinear component reveals a level dependent cochlear pressure reflectance; $|R_e|$ is proportionally greater at frequencies of low power absorption, since $|p_{ae}|$ is inversely proportional to intensity. Therefore, it is concluded that at a stimulus level of 20 dB SPL, subject SS response is predominately linear, more so than for subject JE.
8.8 Modelling stimulus frequency otoacoustic emissions

From equations 8.2 and 8.21, normalising the ear canal sound pressure by the Thevenin open circuit sound pressure, $p_s$, leads to, for $A = A_h$:

$$\frac{p_{ec}[A_h]}{p_s[A_h]} = \frac{AZ_c^o + B}{AZ_c^o + B + Z_s(CZ_c^o + D)} \quad (8.30a)$$

and, for $A = A_L$:

$$\frac{p_{ec}[A_L]}{p_s[A_L]} = \frac{AZ_c^o(1 + R_c) + B(1 - R_c)}{AZ_c^o(1 + R_c) + B(1 - R_c) + Z_s(CZ_c^o(1 + R_c) + D(1 - R_c))} \quad (8.30b)$$

Subtracting equation 8.30a from equation 8.30b gives,

$$\frac{p_{em}[A_L]}{p_s[A_L]} = \frac{2R_cZ_cZ_e^o}{(AZ_c^o + B + CZ_c^o Z_e + DZ_c)(AZ_c^o(1 + R_c) + B(1 - R_c) + CZ_c^o Z_s(1 + R_c) + DZ_s(1 - R_c))}$$

Therefore, the subtracted high from low level normalised ear canal sound pressure (equation 8.31), i.e., the normalised stimulus otoacoustic emission sound pressure, is contingent upon the value of $R_c$. At low levels of stimulation, the origin of auditory microstructure can be identified as an increase in $|R_c|$; i.e., the magnitude of the basal TW relative to the apical TW increases. Through equation 8.31, the existence of OAEs can be viewed as a direct result of changes in $|R_c[\omega, A]|$ with stimulus level.
Specifying the ABCD-transfer coefficients equivalent to a transmission matrix describing a lossless uniform transmission line (equation 8.5) cascaded with an ideal lossless transformer (equation 8.4) leads to:

\[
\begin{bmatrix}
A & B \\
C & D
\end{bmatrix} = \begin{bmatrix}
n \cosh[\Gamma_{ec}/l_{ec}] & Z_{ec}^o/n \sinh[\Gamma_{ec}/l_{ec}] \\
n/Z_{ec}^o \sinh[\Gamma_{ec}/l_{ec}] & 1/n \cosh[\Gamma_{ec}/l_{ec}]
\end{bmatrix} \tag{8.32}
\]

Equation 8.31, together with 8.32, explicitly makes the normalised stimulus frequency otoacoustic emission a function of the object that is being measured, as well as the source impedance of the measurement probe, \(Z_s\). Evaluation of the interaction between \(Z_s\) and \(Z_{ec}\) (section 7.6) showed that power injected from the probe approximated to a constant volume velocity drive condition, in which \(|Z_s| \gg |Z_{ec}|\). A constant volume velocity drive condition is favourable since changes in the load impedance, \(Z_{ec}\), are followed by almost proportionally equal changes in sound pressure; i.e., that attribute of sound which is most readily measurable with the use of microphones.
Figure 8.7: Simulated stimulus frequency otoacoustic emission as a function of frequency. Parameter is cochlear pressure reflectance. Modelled using elementary analogue network of peripheral auditory system (figure 8.2) with standard ear canal parameters (table 8.1) and incorporating Zwislocki middle ear (figure 8.8). Cochlear pressure reflectance against frequency described by vector of constant magnitude, rotating with a spectral periodicity of 1/15 f. Curves plotted at five cochlear pressure reflectance magnitude values; 0.25 to 1.25 in 0.25 steps, annotated a-e.

Referring to figure 6.10, assuming a cochlear pressure reflectance variation with frequency as proposed by Kemp (1980), the resulting SFOAE is described by a vector sweeping out a disc (figure 8.7). However the disc is off-centre, consequently |p_cm| exhibits a microstructure. Such a non-invariant |p_cm| characteristic is observed
8.9 Utility of group delay applied to ear canal acoustic measures

For an ideal network, the phase between a signal that propagates through a network, either lumped or distributive, relative to the input signal, is linear with frequency, resulting in a delay which is constant, and therefore independent of stimulus frequency. A deviation from the ideal linear phase response results in a delay that is a function of frequency, and can be considered as a signal distortion of phase. A measure of the delay between the propagating signal to the input signal for a band of frequencies is given by the parameter group delay, $\tau_\beta [\omega]$ defined as:

$$\tau_\beta [\omega] = -\frac{d(\beta[\omega])}{d\omega}$$  \hspace{1cm} (8.33)

that is, the negative slope of the network phase function (Ruston and Bordogna, 1966).

However, the pressure reflectance is the ratio of the backward to forward travelling pressure waves propagating along a line and, therefore, can be considered as a transfer function. Voss and Allen (1994) define group delay, with reference to the pressure reflectance:

$$\tau_R [\omega] = -\frac{d(\text{Arg}(R_\infty))}{d\omega}$$  \hspace{1cm} (8.34)

i.e., as the negative slope of the pressure reflectance argument. Both $\tau_\beta$ and $\tau_R$ have dimensions of time.

The interpretation of $\tau_\beta [\omega]$ and $\tau_R [\omega]$ initially appears to be rather abstruse, since for positive group delay it follows:
a constraint which for the data given in figures 6.9 and 7.3 does not hold. In fact, referring to a second order bandstop network the negative phase slope constraint patently does not hold for the lossy system (see figure 8.8).

**Figure 8.8:** Network transfer function for a second order bandstop filter. Network takes the form of an inverted-L section with a series and a shunt branch. Series resistance, $R=100 \Omega$. Shunt capacitance $C=10 \mu F$, inductance $0.1 H$. Parameter values chosen to give a notch frequency, $\omega_o=1000 \text{ rad s}^{-1}$. Parameter $R_L$ is the shunt resistance. Left panel is the attenuation function, right panel shows phase function.

To investigate the dependency of $\alpha[\omega]$ on $\beta[\omega]$ for linear minimum-phase systems, such as the second order passive bandstop filter as shown in figure 8.8, it is didactic to refer to the constraint imposed by the Hilbert transform pair (Bode, 1945). Taking the natural logarithm of the exponential function, $\gamma[\omega]$, gives:

\[
\ln[H(\omega)] = \alpha[\omega] + j\beta[\omega]
\]
resulting in one half of the Hilbert transform pair:-

\[ \beta[\omega] = \frac{1}{\pi} \text{PV} \int_{\omega}^{\infty} \frac{\alpha[\omega']}{\omega' - \omega} d\omega' \]  

(8.37)

where PV indicates the Cauchy Principal Value of the integral is computed because of the singularity at \( \omega' = \omega \). The integrand is then a product of the attenuation function and \( (\omega' - \omega)^{-1} \). The operation of the Fourier transform on a sequence relating to an impulse response of a minimum-phase, and therefore causal, system, results in a even real function and odd imaginary function, an example of a Hermitian function.

For both the lossless and lossy case \( \beta[\omega] = 0 \), indicating that the lower integration value, i.e., between \(-\infty\) and \( \omega_o \) is equal but opposite to the upper integration value, i.e., between \( \omega_o \) and \( +\infty \), such that the integrand is a odd function symmetric about \( \omega_o \). However, for the lossless case, a small deviation from \( \omega_o \) results in a large change in the integrand due to the steep gradient in \( \alpha[\omega] \), illustrated by the rapid change in \( \beta[\omega] \). On the other hand for the lossy case, \( \alpha[\omega] \) changes slowly with \( \omega \), such that deviations from \( \omega_o \) result in approximately the same integrand.

Qualitatively, to conclude, the behaviour of the phase function of a minimum-phase system is related to the attenuation function through the Hilbert transform. For highly dissipative systems, i.e., systems manifesting low Q-factors, where the attenuation function is slowly changing, the phase function is consequentially slowly changing. If the attenuation function displays a minimum, the gradient of the phase function becomes positive through the minimum region.

At mid frequencies the occluded ear canal is not lossless because the energy reflectance is low, approximately 0.35 (Keefe et al, 1993); two-thirds of the incident energy is absorbed. The breaking of the negative phase slope constraint (equation 8.35) is seen in data published previously (Kemp and Chum, 1980; Zwicker and Schloth, 1984; Shera
and Zweig, 1993). Hence, when considering dissipative systems such as the ear the concept of group delay loses its utility.

### 8.10 Summary

In this chapter an elementary analogue network model of the peripheral auditory system is presented, enabling cochlear input impedance and pressure reflectance to be inferred from empirical and derived acoustic parameters shown in chapters 6 and 7.

Assuming linear one-dimensional long-wave motion of the $TW^*$, permits a transmission line formalism to be applied. Changes in the wavelength are small as the $TW^*$ propagates through the basal turn, enabling an approximate expression for cochlear input impedance to be found. Inferred and theoretical cochlear input impedances were shown to be of the same order. The role of cochlear pressure reflectance upon model cochlear input impedance was investigated. Comparing inferred to model cochlear input impedance suggested that at a stimulus level of 20 dB SPL cochlear pressure reflectance magnitude was approximately 0.25. Cochlear input impedance microstructure magnitude was shown to be greater than the variation in ear canal input impedance, indicating the role of the occluded ear canal and middle ear upon measurements of auditory microstructure, the origin of which is intracochlear.

A cochlear pressure reflectance characteristic described by a vector of constant magnitude rotating uniformly with linearly increasing frequency, as hypothesised by Kemp (1980), was shown to be consistent, through the use of an elementary model of the peripheral auditory system, with inferred input impedance and simulated stimulus frequency OAEs in the low level linear regime.
Chapter 9: Auditory threshold microstructure

9.1 Introduction

Several workers have observed that psychoacoustic measures, such as absolute pure tone thresholds, low-level loudness contours and pitch, do not vary slowly against frequency, but rather exhibit a rapid undulating characteristic (Elliot, 1958; Van den Brink, 1970; Thomas, 1975; Kemp and Martin, 1976). Furthermore, Kemp (1979a) showed that such an undulating characteristic is commensurate with microstructure manifest in purely acoustic ear canal parameters. Measurements of acoustical parameters displaying auditory microstructure suggest that the mechanics of the peripheral auditory system, in particular the cochlea, contribute, if not completely, then significantly, to such psychoacoustic phenomena, as discussed in chapter 2, Kemp (1979a), Wilson (1980b) and Zwicker and Schloth (1984).

The ear can be stimulated across frequency under a variety of drive conditions. Continuous tonal sound pressure measurements, shown in chapter 6, were elicited by a constant voltage applied onto the sound source which, due to the high acoustical impedance of the electroacoustic probe relative to the occluded ear canal input impedance, approximated to a constant volume velocity source or an iso-volume-velocity drive condition (figure 7.1). Under such a paradigm, neither intra-cochlear absolute nor relative acoustic parameters were known.

When determining absolute threshold to obtain reliable microstructure data, the voltage amplitude applied to the sound source is adjusted at each closely spaced frequency. Consequently, during an absolute threshold paradigm the iso-volume-velocity drive condition is not maintained. The drive condition is then termed iso-sensation and the voltage applied can be thought to be a measure of auditory function, i.e., the ability of a subject to detect a tonal stimulus. Stimulation of the ear under an iso-sensation drive condition will reveal a more fundamental, or functionally relevant, relationship between the various ear canal acoustic parameters defined in previous chapters.
This chapter details the method used to determine absolute auditory threshold, provides evidence of microstructure in auditory threshold and discusses the relationship between specific ear canal parameters and absolute threshold.

9.2 Methods for the determination of absolute auditory threshold

Absolute auditory threshold\(^1\), in common with other psychophysical measures, can be described by a psychometric sigmoidal function, i.e., a graph of the percent positive response against stimulus intensity reveals an S-shaped curve. In the case of a threshold experiment, positive response is read as signal present. The existence of a psychometric function implies that judgements on the presence or absence of a signal are made in noise, where the derivative of the psychometric function describes the probability distribution of the signal in the noise. Absolute threshold can then be arbitrarily defined as a specific point along the psychometric curve. For example, a 50 % point on the ordinate corresponds to \( X_{50} \) on the abscissa, i.e., the signal at the \( X_{50} \) level is detected for half of the total number of presentations.

Fechner (1860) was first to establish a link between sensation, i.e., an attribute of perception\(^2\), and the evoking stimulus, i.e., a physical measure. Hence, the use of the general term psychophysics, or, in the more narrow arena of audition, psychoacoustics. For example, frequency forms the major basis for, and the physical correlate of, the psychoacoustical measure of pitch\(^3\).

Classical psychophysical techniques, as developed by Fechner (1860) and reviewed in Levitt (1971) and Gulick et al (1989), can be categorised into three methods; method of constant stimuli, method of limits and method of adjustment.

---

\(^1\) The object under investigation is a biological system with inherent noise, such that many stimulus presentations are required in order to determine a threshold, implying that a threshold is statistically determined. In addition, threshold is categorised into three types; absolute, terminal and difference.

\(^2\) Perception versus sensation. Analogous to Newton's notion of colour, sound in the environment does not possess pitch, but rather pitch is dependent upon an attribute, most notably frequency, of the impinging sound, the sensation of which adds to our perception.

\(^3\) Sound intensity is a, smaller, factor in pitch perception.
Chapter 9: Auditory threshold microstructure

The method of constant stimuli requires a predetermined set of stimulus levels, usually seven, covering the transition zone from almost never heard to almost always heard, therefore bracketing the absolute limen. Stimuli are presented in a random order allowing the subject to indicate whether the stimulus was detected. After each stimulus is presented a number of times, per cent 'yes' against increasing stimulus level is plotted, revealing an approximate sigmoid curve. The selection of the stimulus level set is crucial in obtaining a good threshold estimate, and it is often necessary to run a preliminary experiment in order to bracket the absolute limen.

Clinical audiometry, an example of method of limits, requires a run of repeated presentations, or trials, of a tone. The intensity of the tone is adjusted as a function of the subject's response. Conventionally, the downward intensity increments are 10 dB while the upward increments are 5 dB. This leads to the observation that many trials of the run do not contribute to the determination of threshold since the intensities are either well below or above the threshold. A second shortcoming relates to bias, which is a function of the choice of initial stimulus level and step size (Levitt, 1971).

Method of adjustment requires the subject to modify stimulus level. The absolute threshold is the average stimulus level at which an audible tone becomes inaudible, and an inaudible tone just becomes audible, repeated a number of times.

The main disadvantage of the classical methods is the role of the subject's response criterion, a measure of the point on the psychometric curve the subject is operating; for example, given two subjects, one may be more reluctant to respond to a barely audible stimulus compared to the second subject. The former subject displays a higher response criterion, making judgements at a high point on the psychometric curve than the second subject, and is therefore deemed, falsely, less sensitive. Methods to minimise, or normalise the response criterion, include the yes-no and forced-choice paradigms, which can be integrated into an up-down method of limits procedure.
Von Békésy (1960) described a new audiometer, the technique of which is now commonly termed Békésy audiometry. The intensity of the pure tone decreases as long as the switch is depressed and increases when the switch is released. The original apparatus allowed the frequency of the pure tone to vary smoothly from 100 to 10k Hz in 15 minutes. It is noted that the original method, for perhaps engineering reasons, describes that the intensity increments were stepped, not varied continuously, in 2 dB steps at a rate of 140 dB per min. The variation in the response gives an indication of the difference limen.

The Audioscan method (Meyerbisch, 1996) differs from the techniques reviewed so far, in as much as, the adjustment of the stimulus intensity is independent of the subject’s response. Typically, the task of the subject is to indicate, by pressing a switch, when the stimulus, a tone of a predetermined intensity is detected as the frequency is swept automatically. The task is repeated over the same frequency region but at ever decreasing intensities. The major disadvantage of the method is that the subject not only is required to make judgements upon the stimulus intensity but, in addition, has to judge the stimulus frequency; a task that appears difficult after a period in which the stimulus is not detected, as would be the case if the subject’s threshold exhibited pronounced microstructure. This additional task effectively increases the cognitive load from that when alternative threshold methods are employed.

Absolute auditory threshold is a function of stimulus duration, the effect being most prominent for duration less than 100 ms. The finding can be interpreted with reference to signal analysis; keeping energy density constant, as the duration of a tone is decreased, it tends towards an impulse, such that its spectrum consists of energy across a greater frequency range. This results in an elevation in threshold due to spectral splattering, in which stimulus energy is delivered to a longer length of the cochlear partition.
9.3 Experimental method

The method chosen to estimate absolute threshold for this study was a sequential adaptive up-down procedure operating at the 50% point. The term sequential adaptive refers to the class of procedures in which the choice of the next stimulus level is dependant upon the subject's past responses. The initial stimulus level was randomly set between 15 and 25 dB SPL, with the initial step size equal to 5 dB. Since a measure of the quality of the estimate is given by the number of turnarounds, reduction in the step size from 5 dB to 2 dB and then 1 dB did not occur at a predetermined trial number, but rather after a specified number of turnarounds; five turnarounds for the 5 to 2 dB transition and a further four for the 2 to 1 dB change. At a stimulus level in dBV corresponding to absolute threshold, given by the mean of the 2 and 1 dB values as illustrated in figure 9.1, the occluded ear canal sound pressure was sampled.

Figure 9.1: Example of subject PS response during an adaptive up-down procedure. Initial stimulus level was randomly set but bounded by 15 and 25 dB SPL. Stimulus level stepped in 5 dB intervals for five turnarounds, then 2 dB for a further four turnarounds, then 1 dB step size. Filled dots refer to stimulus level at each trial. Dashed line is the mean of the 2 dB and 1 dB stepped values. Typical time to obtain 25 trial responses was 37 s.
Each trial consisted of a two interval forced choice paradigm in which two presentations were applied, each 0.5 s in duration. The stimulus, or probe tone, was present in one and absent in the other presentation, the order determined randomly. The task of the subject was to judge for which presentation the probe tone was present, responding via the keyboard. For a correct response the stimulus level was decreased, while for an incorrect response an increase in stimulus level followed.

The stimulus level in dBV corresponding to threshold was given by the mean of the drive voltage values within the 2 dB and 1 dB step regime, as illustrated in figure 9.1. The quality of the threshold value is dependant upon the number of turnarounds per trial, since variance is determined by the deviation from the mean; the greater the number of turnarounds for a given trial number leads to a lower variance, and therefore a more robust threshold estimate.

To obtain threshold estimates across a frequency encompassing a small number of microstructure cycles, the order of stimulus frequencies was randomised to reduce systematic bias in the subject response.

Ear canal sound pressure at threshold was measured by applying a drive voltage corresponding to the threshold estimate and using a lock-in amplifier in quadrature mode. A time constant, typically 1 s, was needed to ensure an adequate signal to noise ratio. Since a duration of at least three times the time constant, equivalent to the settling time of the instrumentation filters, is required to ensure a reliable measurement, each sound pressure measurement took 3 s. The time to acquire a threshold estimate in dBV and dB SPL was approximately 40 s.

---

4 In order to operate at the 71% point of the psychometric curve the responses can be transformed as outlined in Levitt (1971); for a decrease in stimulus level two consecutive correct responses have to be obtained, while an increase in stimulus level followed a incorrect response.
9.4 Experimental Results

Figure 9.2 shows absolute threshold data described in terms of ear canal sound pressure magnitude (dB SPL) for three subjects. It is seen that threshold is non-invariant across each frequency range, varying by up to 15 dB as first reported by Elliot (1958) and subsequently by Van den Brink (1970), Thomas (1975) and Kemp and Martin (1976). The spectral periodicity is discernible for subject KH, approximately equal to 80 Hz, but less so for subjects SS and PS. Hence, the absolute threshold data manifests a microstructure as defined in section 2.1.

Figure 9.2: Measurements of pure tone absolute auditory threshold microstructure. Filled symbols: threshold corresponds to $X_{50}$ level in dB SPL, i.e., $p_{ec}[\omega, X_{50}]$. With reference to section 2.5 and Kim (1983), the difference between the threshold estimate in dB SPL and dBV, i.e., $X_{50}[\text{dB SPL}] - X_{50}[\text{dBV}]$, is plotted as empty symbols. Dashed line: frequency of spontaneous otoacoustic emission (SOAE) of magnitude greater than -15 dB SPL and resolution ±6 Hz. Absolute threshold and SOAEs recorded at same probe placement.
Figure 9.2 shows the frequency of spontaneous otoacoustic emissions (SOAEs) of amplitudes greater than -15 dB SPL, recorded at the same probe location as threshold measurements. Narrowband analysis, between 0 and 6 kHz, of ear canal sound pressure using the ILO88 search mode method is shown figure 6.1. A figure of -15 dB SPL was chosen as a lower limit to ensure good signal to noise in the SOAE. In addition, from figure 9.2 the ear canal sound pressure at threshold is never less than -10 dB SPL, implying that the effect of an SOAE of amplitude less than -15 dB SPL will become negligible.

The frequency of SOAEs of amplitudes greater than -15 dB SPL approximately coincide with the frequency of absolute threshold minima, i.e., auditory sensitivity maxima (figure 9.2). While a statistical relation could not be found for the presence or absence of an SOAE close to threshold minima, Schloth (1983) found that, for a given SOAE, the local minimum in threshold fell within ±10 Hz, concurring with Horst et al (1983), Zwicker and Schloth (1984), Dallmayr (1987) and Long and Tubis (1988). Such a correspondence between the frequency location of an SOAE and threshold minima could arise through the interaction of the applied stimulus tone and the SOAE resulting in the phenomenon of beats, the presence of which may aid detection, as suggested by Wilson (1983).

Kemp (1979a) reported that, across a frequency region bracketing a threshold minimum the quality of a single frequency stimulus appeared 'blurred', indicating the presence of a second, non-stimulus tone, presumably an SOAE, interacting with the stimulus, resulting in the beat phenomenon. Since the region of 'blurred' quality was bounded at the lower level by the threshold curve, Kemp (figure 7b, 1979a) designated the synchronisation threshold to provide the upper boundary, defined as the minimum stimulus level required to convey a clear percept of pitch. At stimulus levels above synchronisation threshold pitch was perceived as 'clear' and the beat phenomenon disappeared.
9.5 Role of SOAE in the calculation of impedance

When using a Thevenin equivalent calibrated probe method to calculate acoustic input impedance as described in section 2.13 it is assumed that the ear canal sound pressure is stable and synchronised to the voltage applied to the sound source, such that the system has attained steady state. Kim (1983) noted, at stimulus levels about absolute threshold the ear canal sound pressure may consist of the sum of three components; the applied stimulus, an SFOAE and an SOAE. The first component is the transformation of the voltage applied to sound pressure developed in the load effected by the sound source. The latter two components are of cochlear origin, but have different consequences upon the calculation of impedance when employing the Thevenin equivalent calibrated probe method.

The effect of an SFOAE, elicited by an evoking pure tone stimulus, can be thought of as a change in the ear canal sound pressure arising through a modification in the ear canal impedance, as shown in figure 7.1. Consequently, for a given voltage amplitude applied to the sound source the method to solve for ear canal input impedance through the potential divider equation (equation 2.1) in the presence of an SFOAE is valid.

An SOAE, by definition requires no evoking stimulus. However, Kemp (1979a) found that the behaviour of a SOAE when a single frequency tone is applied falls into two states, dependent upon the level and frequency of the applied tone relative to the SOAE. Either the resultant ear canal sound pressure is stable and phase locked, and its spectrum is single peak in which the SOAE is synchronised to the applied stimulus, or two distinct spectral peaks exist, such that the SOAE is non-synchronised. Thus, the Thevenin equivalent calibrated probe method to calculate acoustic input impedance remains valid when a synchronised SOAE is observed since its effect is identical to that of a SFOAE. In contrast, the presence of an non-synchronised SOAE breaks the synchronicity between dBV and dB SPL preventing the input impedance to be calculated through equation 2.1. The degree of the deleterious effect of an non-synchronised SOAE upon calculation of impedance is dependent upon the frequency separation between the non-synchronised SOAE and stimulus frequency relative to the
analysis bandwidth governed by the lock-in amplifier time constant. For example, when the non-synchronised SOAE falls within the measurement bandwidth the resultant ear canal sound pressure becomes unstable and desynchronised to the applied voltage. The method described to calculate input impedance then becomes invalid.

In the experiment described, assessment of the synchronicity effect of an SOAE by either a subjective measure of pitch clarity or ear canal acoustic temporal analysis was not undertaken. It is for this reason that threshold data in the presence of SOAE is rejected for further analysis.

When calculating ear canal acoustic input impedance with the Thevenin equivalent calibrated probe method, the issue of a non-synchronous SOAE can be overcome by determining the synchronisation threshold, rather than establishing the absolute threshold (Kemp, 1979a; Wilson and Sutton, 1981; Zwicker and Schloth, 1984). The synchronous threshold demands a modification in the experimental psychoacoustic paradigm in which a judgement of pitch clarity is determined.

9.6 Quantifying absolute threshold of sensory systems

Living organisms possess sensory systems that convert, or transform, many modes of energy that exist in Nature. In humans, the eye, vestibular system and cochlea, through specialised adaptation, detect energy in the mode of electromagnetic visible light, rotational acceleration and sound, respectively. The detection mechanism for each sensory system attains a high degree of sensitivity, as well as a narrow selectivity and wide dynamic range. Khanna and Sherrick (1979) note that the practical measurement unit of absolute threshold of a sensory system normally takes the stimulus dimension that is most readily quantified by the use of conventional measurement devices.

Absolute auditory threshold can be defined by one of a number of parameters; for example, minimum audible field, dBV applied to the sound source, ear canal dB SPL or power absorbed at the TM. In clinical audiometry, a standardised reference level in dB HL is used, equivalent to the normalised dB SPL for normal average hearing. However, during clinical audiometry the SPL developed in the ear is not sampled directly, but
inferred from knowledge of the transfer function of the sound source operating into an artificial ear. In the case of the auditory system, absolute threshold is traceable to relative units of sound pressure, due to the use of reliable stable sensitive pressure microphones.

Sound pressure as a descriptor of sound is deficient because its magnitude tends to zero at particular points within systems that exhibit interference, a wave phenomenon arising through phasic interaction of waves travelling in different directions. However, energy in a system is a more fundamental physical quantity, defined as the capacity, or ability, of that system to do work. There is no fixed relationship between energy and sound pressure. To determine energy absorbed in a load requires knowledge of the sound pressure developed across the load and its impedance. Energy lends itself to the description of transient stimuli, whereas for a continuously applied signal, power is a more appropriate stimulus descriptor.

Assuming one dimensional motion, absolute power transmitted to the middle ear is a function of absolute power incident and non-dimensional transmittance. Approximating the ear canal to a lossless transmission line, i.e., a line that has zero capacity to absorb power, the absolute power incident upon the TM, $\Pi_{ec}'$, is made with reference to the ear canal impedance, $Z_{ec}$, in particular the real part, and the ear canal sound pressure, $p_{ec}':$

$$\Pi_{ec}' = \frac{|P_{ec}'|^2}{\text{Re}\{Z_{ec}\}}$$

(9.1)

The finding that the cochlea supports a backward travelling wave, delayed relative to the forward travelling wave (Kemp, 1978), which due to the coupling afforded by the ME, modifies the ear canal input impedance, implies that the coupling of the TM to the ear canal, and therefore the power transmittance, $T_{ec}$, varies against frequency. Power transmittance is defined as:

$$\Pi_{ec}^\theta = T_{ec} \Pi_{ec}'$$

(9.2)
In the following section, power flow into the ear canal, designated $\Pi_{ee}$, will be used as a measure of absolute threshold.

### 9.7 Ear canal power flow under an iso-sensation drive condition

By calculating ear canal input impedance (equation 2.1) and estimating ear canal reflectance (equation 7.3), the power constituents were obtained; incident and absorbed ear canal power shown in figure 9.3. With increasing frequency the power incident undulates by up to 10 dB. The intensity level reference, standardised to $10^{-12}$ Wm$^{-2}$, approximates to the average threshold intensity for mid frequencies in humans. For an ear canal radius equal to 3.76 mm (ANSI S3.25-1979), the power incident upon the TM is then equal to -13.5 dB re 1 fW. The ear canal power microstructure measured for subject PS waxes and wanes about this reference threshold intensity level.

Ravicz and Rosowski (1994) measured acoustic power flow in the ear of the Mongolian gerbil. In contrast to the power absorbed microstructure shown in figure 9.3, Ravicz and Rosowski (1994) report that, at a stimulus level corresponding to absolute threshold, acoustic power absorbed at the TM was constant across frequency, equal to $10^{-16}$ W (-10 dB re 1 fW). Since the study is reported in the form of a conference proceeding abstract, the experimental method and data analysis is only sketched. However, it has been stated that when comparing rodent and human ears there exists a significant difference in the anatomical features of the cochlea, in particular the organ of Corti, that play a crucial role in the degree of auditory microstructure. Assessing anatomical data on human and guinea pig cochleae presented in Wright (1984), Shera and Zweig (1993) conclude that the organ of Corti displays an almost crystalline regularity in the guinea pig whereas, in the human the cellular arrangement is more disordered. Since cellular disorder along the organ of Corti suggests mechanical impedance irregularities, the degree of scattering of the TW$^+$ in the smooth rodent cochlea would be less than that in the human cochlea.

By using scanning electron microscopy, Wright (1984) qualitatively showed that the surface of the organ of Corti in guinea pig is more ordered than the human cochlea.
Although the human cochlea was excised from a 50 year old male who prior to death displayed a normal audiogram when normalised for age and sex, caution must be exercised when making inter-species comparisons because of the small sample set of material used.

**Figure 9.3:** Iso-sensation ear canal power flow. *Incident power separated into reflected and absorbed for subject PS. Absolute threshold determined using two interval forced choice sequential adaptive up-down paradigm and shown in figure 9.2. Power in dB re 1 fW (10^{-15}).*

When incident and absorbed power is plotted on a logarithmic scale the distance between the two curves is proportional to power transmittance. From figure 9.3 it is seen that power transmittance varies across frequency; at frequencies corresponding to intensity maxima, i.e., low sensitivity, transmittance is low, whereas at intensity minima transmittance increases.
Chapter 9: Auditory threshold microstructure

Under an iso-volume-velocity stimulus within the low level linear regime the incident ear canal power undulates against frequency, as shown in figure 7.4. Such a finding implies that the coupling between the source and ear varies against frequency. Consequently, the ear at particular frequencies extracts more power from the source than at other frequencies.

The mechanism effecting a modification in coupling between the source and ear, i.e., the presence of a TW⁻, exists under a low-level iso-sensation drive condition. An undulation in incident power may contribute to, or to a greater extent wholly explain, absolute threshold microstructure. In order to assess the issue of coupling and its role in threshold microstructure a simple phenomenological model of power flow is presented below.

9.8 Phenomenological model of power flow in the peripheral auditory system

Under an iso-sensation drive condition the power flowing past the detector, \( D \), is assumed to be constant, i.e., detector efficiency is constant against frequency. Assuming that the power dissipated by the ME is small relative to the power transmitted to the cochlea, figure 9.3 approximates to the power flow constituents at the OW⁵. Since the absolute input power and relative output power are known, it follows that the relative, but not absolute, power characteristic within the peripheral auditory system is solvable:-

\[
\Pi_i \times T_{ec} \times E = D
\]

(9.3)

Delivery efficiency, \( E \), was estimated given the empirically determined power incident, \( \Pi_i \), and transmittance, \( T_{ec} \). The parameter \( E \) can be thought of as the efficiency with which the available, absorbed power is delivered to the detector.

---

⁵ Dissipative losses in the middle ear are most likely to be at the incudo-malleus joint, incudo-stapedial joint (Zwislocki, 1962).
Chapter 9: Auditory threshold microstructure

Figure 9.4: a) Phenomenological model of power flow in the peripheral auditory system. b) Power flow model parameters. Power incident and transmittance derived from iso-sensation experiment for subject PS. Power gain calculated assuming $D$ invariant with frequency.

a)
Delivery efficiency

Transmittance

Power incident

Frequency (Hz)
From the model of power flow presented in figure 9.4a, threshold microstructure arising, wholly, from a variation in the coupling of the ear to the sound source would result in a delivery efficiency parameter value invariant with frequency. The finding that $E$ undulates by up to 8 dB implies that the power absorbed at the TM is delivered to the detector with varying efficiency (figure 9.4b).

Initially, this appears a remarkable finding. However, the cochlea acts, to a first approximation, as a one dimensional inhomogeneous distributed transmission line supporting a forward ($TW^+$) and, at low level stimulation, a backward ($TW^-$) travelling wave. The $TW^+$ and $TW^-$ interact, or more precisely sum vectorially, leading to constructive or destructive interference at particular stimulus frequencies. The consequence of wave interference within the cochlea upon the efficiency with which the available power is delivered to the detector is illustrated in the following section.

9.9 Illustration of standing waves within the cochlea

Throughout this thesis, discussion on the phenomenon of auditory microstructure has been made with reference to the existence of a basalward cochlear travelling wave ($TW^-$). In order to illustrate the effect of the $TW^-$ upon the distribution of energy along the length of the cochlea a heuristic model is implemented. Since the aim of the model is illustrative it in no way constitutes a numerically rigorous attempt to model the response of a real, physiologically healthy cochlea.

9.9.1 Bi-directional energy flow model

The hydro-mechanical characteristic of the cochlea is modelled as an inhomogeneous one dimensional linear transmission line. The formulation is only sketched here as it is detailed elsewhere, notably in de Boer (1984). Two immittance parameters completely characterise the hydro-mechanical property of the transmission line at each point; the series impedance per unit length, $Z_F$, and shunt admittance per unit length, $Y_{CP}$ (figure
9.5). The acoustic property of the cochlear fluid is carried in $Z_F$, while $Y_{CP}$ describes the cochlear partition admittance.

Such a transmission line formalism is equivalent to the long-wave approximation in that the wavelength of the travelling wave is greater than the height of the scalae. This condition is observed for travelling waves launched under low to mid frequency stimulation of the OW and propagating through the basal turn (de Boer, 1980). The condition breaks down as the TW$^+$, approaching a place characteristic for the stimulus frequency, slows down. Consequently, its wavelength decreases and the long-wave approximation becomes invalid about the TW$^+$ peak. Although quantitatively inexact, the long-wave formulation is adequate for the purposes of illustration because the qualitative features of wave motion about the TW$^+$ peak are captured.

**Figure 9.5:** Section of cochlear duct, of length $\Delta x$, modelled as a one dimensional linear transmission line. Each section characterised by two immittance parameters, $Z_F$ and $Y_{CP}$, the series impedance per unit length and shunt admittance per unit length, respectively. $p_e$ is the trans-partition sound pressure. Volume velocity of cochlear fluid is designated by $U_c$. Complex power is given by $p_e(U_c)^*$ where the asterisk denotes complex conjugation.
The energy per unit length of the apical travelling wave, $E^+$, was determined by the following method. By obtaining the attenuation constant, given by the real part of the propagation constant, the power flowing into each section, $W$, was calculated. Group velocity, $U$, was determined using the method given in Lighthill (figure 61, 1978). Energy per unit length, $E^+$, equal to the ratio $W/U$, was found.

Assuming that the dissipation between the base and apical reflector is small, the nature of the energy per unit length of the TW$^-$ was set equal to $E^+$ for the following reason. Consider a lossless transmission line under harmonic stimulation with an ill-matched terminating impedance. A snapshot in time will reveal the backward travelling wave identical in shape, to the forward travelling wave, differing only in amplitude and phase dependent upon the terminating impedance relative to the characteristic impedance.

At stimulus levels commensurate with absolute threshold a linear behaviour in auditory parameters is observed (Littlefield, 1973; Russell and Sellick, 1978; Robles et al, 1986; Zwicker and Schloth, 1984; Shera and Zweig, 1993). Since the principle of superposition is valid for all linear systems, combining the forward and backward travelling waves is equivalent to vectorial addition. The energy distribution of the TW$^-$ was set to a fractional value of $E^+$. The total energy per unit length is given by a vectorial sum of $E^+$ and $E^-$ where the phase difference is equal to the phase of the $p^-$ relative to $p^+$, calculated using the WKBJ approximation (Zweig et al, 1976). The phasic interaction of the forward and backward travelling waves leads to constructive and destructive interference at particular points from the base. Such interaction results in standing waves in the energy distribution (figure 9.6).
**Figure 9.6:** Illustration of standing waves within the cochlea. The total energy per unit length against distance from base is shown for two frequencies; 1000 and 1040 Hz. Frequencies chosen to give a maximal pressure reflectance difference argument, i.e. \( \angle R_r[1000 \text{ Hz}] - \angle R_r[1040 \text{ Hz}] = \pi \), when \( \Delta f / f \approx 1/15 \) (Zweig and Shera, 1995). Pressure reflectance magnitude at the base equal to 0.3. The energy at the base is constant for both frequencies, i.e. iso-input energy drive condition. The mechanism relating to the reflection of TW\(^{-}\) is therefore not made explicit, rather the pressure reflectance at the base is specified.

For convenience, an iso-input energy drive condition is specified, such that the energy developed at the base is constant against frequency. The role of the TW\(^{-}\) in threshold microstructure is then illustrated as a maximal peak in energy that is frequency dependent; a pattern seen in figure 9.6. For a constant detector efficiency, the undulation in the efficiency parameter \( E \) of the power flow model (figure 9.4) can be interpreted as arising from a rapid change in maximal energy across the cochlear partition against frequency.
9.10 Summary

This chapter considered aural stimulation under iso-volume-velocity and iso-sensation drive conditions in relation to functional relevance. Methods for determining psychophysical measures, in particular absolute auditory threshold, were discussed. The role of synchronous and non-synchronous otoacoustic emissions in the measurement of input impedance using the Thevenin equivalent probe method was discussed. The invalidity of the method in the presence of a non-synchronised spontaneous OAE is surmounted by stimulating the ear at the synchronous, rather than absolute, threshold.

It was argued that characterising absolute threshold in terms of energy, rather than sound pressure, is more meaningful since it quantifies the capacity to do work. Auditory microstructure was evident in absolute threshold and power absorbed experimental data. Modification in coupling between the ear and sound source across frequency was assessed with the use of a model of power flow within the peripheral auditory system. It was concluded that threshold microstructure was partially, rather than wholly, attributable to changes in coupling against frequency.

An inhomogeneous transmission line model of cochlear dynamics was used to illustrate the effect of the basally travelling wave and its role in standing waves. The maximal energy in region of the peak of the travelling wave changed greatly for changes in frequency corresponding to half the microstructure spectral period.
Chapter 10: Conclusions and further research

10.1 Conclusions

The primary aim of this thesis was to propose the hypothesis that auditory microstructure is a consequence of bi-directional energy flow within the cochlea. It is argued that the presence of basalward travelling energy modifies cochlear input impedance from that case when energy propagates only apically. The modifications were observed experimentally in ear canal acoustics and accounted for with the use of an analogue network model of the peripheral auditory system. The following conclusions have been drawn:

1) By providing the theory of the least mean squared method to calculate model parameters of an overdetermined system the connection between the system equation configuration and solution matrix was demonstrated. Amelioration of the least mean squared method applied to the Thevenin calibration of an electroacoustic probe was developed allowing for the estimation, and minimisation, of standard errors in the model parameters. The equivalent Thevenin source parameters of the electroacoustic source were presented. It is concluded that the method employed has greater utility than the previously published methods because the uncertainty in the model parameters can be computed.

2) Evanescent, as well as propagating modes are launched when sound energy is injected into the ear canal. The role of the evanescent wave in the measurement of input impedance was investigated. The magnitude of the evanescent wave relative to the plane wave was shown to be greatest at frequencies corresponding to minima in input impedance magnitude. The effect of the evanescent wave upon the input impedance measurement was shown experimentally. An analysis of the propagation of errors through the impedance measurement, arising from both the calibration of the probe and measurement of sound pressure, was expressed in order to quantify the uncertainty. In order to verify the calibration, the acoustic input impedance of a hard walled rigidly terminated tube was measured using the Thevenin calibrated probe method. The measured data was shown to be in good quantitative agreement to model predictions.
when the errors were taken into account. In conclusion, the effect of the evanescent wave upon measurements made within the human ear canal is smaller than in highly resonant loads because the ear presents a greater resistance.

3) It was hypothesised that ear canal sound pressure measurements are a function of the acoustical characteristic of the probe assembly. Measurement in human ear canals of continuous tonal aural sound pressure (CTASP) against stimulus level displayed a nonlinear characteristic. Microstructure in CTASP, as a function of frequency, was observable at a stimulus level of 20 dB SPL, but undetectable at 60 dB SPL. A lumped parameter network model of ear canal acoustics was implemented. From network analysis it is concluded that such measurements carry information on the source impedance, which must be considered when making and interpreting ear canal acoustic measurements. For example, data on spontaneous otoacoustic emissions suggests that a relationship exists between the presence of an SOAE and probe reflectance. The vector-subtraction method used to reveal a nonlinear component in ear canal sound pressure, was performed. The error in the vector-subtracted signal was expressed, indicating the role of gaussian distributed noise during the measurement and the accuracy in the attenuation of the stimulus.

4) By utilising the lumped parameter network model, it was hypothesised that, in addition to sound pressure, acoustic parameters such as impedance, reflectance and power exhibit a microstructure characteristic. Acoustic input impedance, reflectance and power flow constituents, derived from ear canal sound pressure data under an iso-volume velocity drive condition, exhibited microstructure at a 20 dB SPL stimulus level, indicating that at particular frequencies the ear extracted greater power than at other frequencies.

5) An elementary analogue network model of the peripheral auditory system incorporating scattering parameters permitted cochlear pressure reflectance to be inferred from empirical data. The inferred cochlear pressure reflectance was non-zero, implying that energy within the cochlea flows basalward as well as apically. The inferred microstructure magnitude in the cochlear pressure reflectance was 4 dB greater than ear canal pressure reflectance, due to decoupling of the probe from the tympanic
membrane. Ear canal acoustic input impedance was modelled through the use of the analogue network and a cochlear pressure reflectance described by a vector of constant magnitude and uniformly rotating with linearly increasing frequency. The predicted ear canal microstructure manifested characteristics observed in the empirical data.

6) As a result of microstructure observed in ear canal input impedance, it was proposed that auditory threshold is determined by the degree of coupling between the ear and sound source. An adaptive up-down psychoacoustic paradigm was employed to determine absolute auditory threshold of subjects used in the previous experiment. Under an iso-sensation drive condition the absolute power absorbed at the tympanic membrane was determined, which assuming negligible dissipation in the middle ear, constitutes the power absorbed by the cochlea. Both power absorbed and power transmittance manifested a microstructure characteristic. The variation in coupling between the ear and sound source partially, but not wholly, accounted for auditory threshold microstructure. A phenomenological model of power flow within the auditory periphery revealed that the efficiency with which the available power is delivered to the detector is non-invariant across frequency. It was argued that such a finding is indicative of systems supporting standing waves, a point illustrated through the use of a transmission line model of energy flow along the cochlea.

From empirical data and model predictions presented in this thesis it is concluded that auditory microstructure arises as a consequence of bi-directional energy flow within the cochlea.

10.2 Further research

The experimental technique reported in this thesis can be implemented into a number of conventional auditory measurements in order to a) provide information on the quality of recording, b) allow an assessment of possible confounding effects and c) permit a more functionally relevant measure of the peripheral auditory system.

a) Neonatal auditory screening programmes employing evoked OAEs typically result in a high false positive rate, i.e. a low figure of specificity, primarily due to abnormal
conductive properties of the occluded ear canal and middle ear. Since the magnitude of an evoked OAE is dependent upon the power absorbed by the occluded ear, absolute power absorbed, $\Pi_{ec}^*$, may be used as a measure of the quality of the OAE recording, by providing information on the status of the signal pathway to the cochlea. Such information may prove useful during the referral procedure.

b) Clinical tympanometry determines the impedance at the tympanic membrane, $Z_{tm}$, at a frequency of 226 Hz. From a functional perspective, ear canal transmittance, $T_{ec}$, rather than $Z_{tm}$, determines the power transmitted to the middle ear for a given power incident upon the tympanic membrane. The Thevenin calibrated probe technique allows for rapid wideband measurement of $T_{ec}$, and therefore a more functionally relevant measure of audition.

c) A modification to evoked OAEs during contralateral stimulation (CS) is often attributed to the activation of the auditory efferent system. However, activation of the contralateral acoustic stapedial reflex may also occur, confounding the interpretation of CS on OAEs. Measurement of $T_{ec}$ in the presence and absence of CS will permit an assessment of the role of the contralateral acoustic stapedial reflex on modifications to evoked OAEs.
Chapter 11: References


Brüel and Kjær (1990) "Instruction manual: Probe microphone Type 4182."


page 203


Turin (1997) Personal communication.


Chapter 12: Appendix: A note on software

Two software packages were extensively used throughout this study; Borland Pascal version 7 for DOS and Mathematica (Wolfram Research) version 2 for DOS.

To calculate the Thevenin parameters of an electroacoustic probe, as detailed in chapter 5, the following procedure was used. On a Brüel and Kjær dual channel analyser Type 2035, magnitude and phase spectra data in comma-delimited format were dumped to a 3.5" floppy disk. On a PC, a pascal program read the data, performed the error minimisation routine to calculate the Thevenin parameters and computed the associated errors. Pascal was chosen over Mathematica because of its increased speed of computation. However, pascal does not support a complex variable, therefore specific functions were written to perform complex arithmetic using real variables of double precision.

A pascal program was written to control an EG&G lock-in amplifier via the serial port using RS232 interconnection, and acquire data during CTASP (chapter 6) and iso-sensation (chapter 9) measurements.

For modelling purposes, Mathematica programs were written to obtain derived and inferred ear canal parameters (chapters 7 and 8) and implement the model of power flow in the auditory peripheral (chapter 9). Mathematica was chosen over pascal for modelling because of convenient graphical routines and complex arithmetic capabilities.