UNIVERSITY OF SOUTHAMPTON

Testing a Model of Stimulus Frequency Otoacoustic Emissions in Humans

Benjamin Lineton

A thesis submitted for the degree of

Doctor of Philosophy

Faculty of Engineering and Applied Science

Institute of Sound and Vibration Research

October 2001

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING AND APPLIED SCIENCE
Institute of Sound and Vibration Research

Doctor of Philosophy

Testing a Model of Stimulus Frequency Otoacoustic Emissions in Humans

by Benjamin Lineton

In recent theories of cochlear mechanics, it has been proposed that spontaneous, transiently evoked and stimulus frequency otoacoustic emissions originate from scattering of a "tall and broad" travelling wave (TW) with a random spatial distribution of cochlear inhomogeneities. In contrast, in an earlier theory, it was proposed that scattering arises from inhomogeneities in the form of regular spatial corrugations. Both these theories successfully predict the existence of quasi-periodic frequency variations in the spectra of these otoacoustic emissions. However, they invoke very different cochlear mechanical properties to explain the frequency spacing commonly characterised by a parameter known as its periodicity. In the first theory, the periodicity is determined predominantly by the wavelength of the TW near its peak amplitude, whilst in the second it is determined by the spatial period of the corrugations, and is therefore largely independent of TW wavelength. The aim of this thesis is to test these two rival theories in humans by attempting to induce changes in the TW wavelength through ipsilateral acoustic suppression, whilst measuring any accompanying changes in the periodicity of stimulus frequency otoacoustic emissions (SFOAE).

For both theories, a one-dimensional longwave model of cochlear mechanics was developed including representations of the two scattering mechanisms and of nonlinear cochlear active processes. Detailed predictions of the changes in SFOAE periodicity under conditions of self-suppression and high-side, two-tone suppression were then made from both models. In the model with random inhomogeneities, the periodicity clearly increased with the extent of self-suppression, and decreased with the extent of two-tone suppression. In sharp contrast, in the model with regular spatial corrugations, no change in periodicity occurred in either case. This result provides a means for differentiating between the two theories experimentally. Experiments were performed in 20 human subjects with normal hearing to measure any changes in SFOAE periodicity during self-suppression and two-tone suppression.

The experimental results were in broad agreement with the theoretical predictions of the model with random irregularities: SFOAE periodicity generally increased with the extent of self-suppression and decreased with the extent of high-side, two-tone suppression. This result was interpreted as strong evidence favouring the theory of spatially random inhomogeneities over the rival theory of regular spatial corrugations, thereby enhancing understanding of the mechanism for generation of otoacoustic emissions.

ACKNOWLEDGEMENTS

First I would like to thank my supervisor, Professor Mark Lutman, for his advice and support over the course of this project. I would also like to thank Professor Phil Nelson and Professor Bob Allen who reviewed the project and who offered helpful comments and suggestions along the way; Amanda Hall who agreed to sit still for hours on end with a probe poking into her ear; Torsten Marquardt, Dr. Sarosh Kapadia and Dr. Matthew Wright with whom I had many useful and interesting discussions; the many members of the Hearing and Balance Centre who made my time there so enjoyable; all the volunteers who participated in my experiments and the University of Southampton which funded my study. Finally, I would like to thank Anna for her help, support and patience.

CONTENTS

PART IV: CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE WORK

LIST OF FIGURES

LIST OF TABLES

LIST OF ABREVIATIONS

AR	Autoregressive
ARMA	Autoregressive, moving average
BM	Basilar membrane
CP	Cochlear Partition
DPOAE	Distortion product otoacoustic emission
IHC	Inner hair cell
MA	Moving average
OAE	Otoacoustic emission
OHC	Outer hair cell
RMS	Root mean square
SFOAE	Stimulus frequency otoacoustic emission
SNR	Signal-to-noise ratio
SOAE	Spontaneous otoacoustic emission
S/EOAE	Spontaneous-to-evoked otoacoustic emission (power rating)
TEOAE	Transient evoked otoacoustic emission
TW	Travelling wave
WKB	Wentzel-Kramers-Brillouin

LIST OF SYMBOLS

c_{TW}	TW phase speed [ms^{-1}].
d	Place-frequency length parameter [m].
$E[...]$	Expectation operator.
f	Frequency [Hz].
f_{GMF}	Geometric mean frequency [Hz] of two frequencies.
$F\{\ \}$	Fourier transform operator.
$F^{-1}\{\ \}$	Inverse Fourier transform operator.
$\boldsymbol{g}(f)$	Frequency response function characterising transmission from earphone to cochlea.
g_0	Specific (frequency independent) value of $\boldsymbol{g}$.
$\boldsymbol{h}(\eta)$	Complex version of the filter impulse response function in the η-domain.
i	$\sqrt{-1}$
$\boldsymbol{k}(f)$	Wavenumber [m^{-1}].
$\mathsf{K}_{xx}(\eta')$	Normalised autocorrelation function of an arbitrary signal (or random process), $x(\eta)$ in the η-domain.
$\tilde{\mathsf{K}}_{xx}(\eta')$	Estimate of the normalised autocorrelation function $\mathsf{K}_{xx}(\eta')$.
L_ε	Logarithmic measure of the poorness-of fit.
$m_0(x)$	CP mass per unit area [kgm^{-2}].
$p(x,y,t)$	Fluid pressure in a 2D model at location (x,y) and time t [Pa].
p_d	Semi difference pressure (half the difference between the fluid pressure in the upper and the lower channel) [Pa].
$\boldsymbol{p}_{EC}(f)$	Ear canal pressure [Pa].
$\boldsymbol{p}_{EC:R=0}(f)$	Ear canal pressure when the cochlear is reflectionless [Pa].
p_s	semi sum pressure (half the sum of the fluid pressure in the upper and lower channels) [Pa].
p_{St}	Fluid pressure at the stapes [Pa].
$\boldsymbol{p}_{SF}(f)$	SFOAE pressure in the ear canal [Pa].
p_{RW}	Fluid pressure at the round window [Pa].
$\boldsymbol{p}_\Delta(f)$	Normalised SFOAE pressure.

$p_{\Delta:\,\mathrm{RMS}}$	RMS normalised SFOAE pressure, calculated by averaging in the η-domain.
$\boldsymbol{Q}_{Src}\,(f)$	Source (i.e. earphone) 'short-circuit' volume velocity [m^3s^{-1}].
$\boldsymbol{r}\,(f)$	Reflectance looking backwards, defined as the transfer function given by the ratio of reflected to incident pressure signals at a given point.
r_0	Specific (frequency independent) value of $\boldsymbol{r}$.
$\boldsymbol{R}\,(f)$	Reflectance looking forwards, defined as the transfer function given by the ratio of reflected to incident pressure signals at a given point.
$\mathsf{R}_{xx}(\eta')$	The autocorrelation function of an arbitrary signal (or random process), $x\,(\,\eta\,)$ in the η-domain.
$\tilde{\mathsf{R}}_{xx}(\eta')$	Estimate of the autocorrelation function, $\mathsf{R}_{xx}(\eta')$.
SNR_1	Estimate of SNR based on the noise in adjacent frequency bands.
SNR_2	Estimate of SNR based on the inter-session variability.
SNR_3	Estimate of SNR based on the analyticity of the SFOAE.
t	time [s].
$\boldsymbol{T}\ (\Theta)$	Travelling wave function for an approximately scaling symmetric cochlea, defined as the velocity response of the CP, normalised with the stapes velocity.
$\boldsymbol{T}_{\mathrm{ij}}\,(f)$	2-port matrix elements for front end model.
$u(x,y,t)$	Fluid x-velocity component in a 2D model at location (x,y) and time t [ms^{-1}].
$u_{ST}(y,t)$	Stapes x-velocity component in a 2D model at location y and time t [ms^{-1}].
$u_{RW}(y,t)$	Round window x-velocity component in a 2D model at location y and time t [ms^{-1}].
$v(x,y,t)$	Fluid y-velocity component in a 2D model at location (x,y) and time t [ms^{-1}].
$v_{CP}(x,t)$	Cochlear partition z-velocity in a 1D or 2D model at location x and time t [ms^{-1}].
x	Longitudinal coordinate along the BM from the stapes [m].
y	Vertical coordinate in the cochlear measured from the BM [m].
$\boldsymbol{Y}_{Src}\,(f)$	Source internal acoustic admittance [$m^3s^{-1}Pa^{-1}$].

z	Lateral coordinate in the cochlear measured from the centre of the BM [m].
$\mathbf{Z}(f)$	Acoustic impedance defined as the transfer function given by the ratio of acoustic pressure to volume velocity [Pa.m^{-3}.s].
$\mathbf{Z}_{CA}(x,f)$	Cochlear amplifier component of the CP point impedance [kgm^{-2}s^{-1}].
$\mathbf{Z}_{CP}(x,f)$	CP point impedance [kgm^{-2}s^{-1}].
$\mathbf{Z}_{Pass}(x,f)$	Passive component of the CP point impedance [kgm^{-2}s^{-1}].
$\mathbf{Z}_{Sc}(x,f)$	Scattering component of the CP point impedance [kgm^{-2}s^{-1}].
α	The α-parameter in the 4-parameter model, which gives a measure of the influence of multiple reflections on the fitted SFOAE ϕ-spectrum.
β	The β-parameter in the 4-parameter model, which gives a measure of the strength of SFOAE.
$\beta_{x\text{-}\omega}$	Ratio of stimulus frequency to the CP natural frequency at a point.
δ_0	CP damping parameter.
Δf	Ripple frequency interval defined as the frequency spacing between adjacent peaks in a ripple pattern [Hz].
$\Delta\eta$	Natural logarithmic frequency ripple interval, defined as the difference between two values of η evaluated at adjacent peaks in a ripple pattern.
$\Delta\eta_{Oct}$	Octave ripple interval, defined as the difference between two values of η_{Oct} evaluated at adjacent peaks in a ripple pattern.
ε_{MS}	Measure of the poorness-of-fit between the measured and fitted ϕ-spectrum.
η	Natural logarithmic frequency scale, defined by $\log_e(f/f_0)$, where f_0 is the highest characteristic frequency in the cochlea.
η'	η lag variable.
η_{Oct}	Octave logarithmic frequency scale, defined by $\log_2(f/f_{Ref})$, where f_{Ref} is an arbitrary reference frequency.
$\theta(f)$	Phase angle of a spectrum, as a function of frequency [radians].
$\Theta(x,\omega)$	$\log_e \beta_{x\text{-}\omega}$
$\hat{\lambda}$	Wavelength at the peak of the travelling wave [m].

ρ	Fluid density [kgm^{-3}].
$\rho_{Sc}(\eta)$	Scattering potential in the η-domain, related to the scattering impedance.
σ_{xx}	Standard deviation of a an arbitrary signal (or random process), $x(\eta)$ in the η-domain.
$\tau_{GD}(f)$	Group delay of a signal or transfer function, defined by the slope of the phase of the spectrum [s].
ϕ	Fourier conjugate variable to η [cycles].
ϕ_{BW}	The ϕ-bandwidth, defined as the width of the peak of the fitted ϕ-spectrum.
ϕ_C	The ϕ-centre value, defined as the value of ϕ at the peak of the fitted ϕ-spectrum. For an idealised SFOAE this is equal to the reciprocal of the periodicity, Ψ.
$\phi_{GCD}(f)$	Group cycle delay of a signal or transfer function, defined as the product $f\,\tau_{GD}(f)$ [cycles].
Ψ	Periodicity, defined by $\Delta f / f_{GMF}$ either for a pair of ripple peaks, or averaged over a series of ripple peaks.
ω	Stimulus radian frequency [rad/s].
ω_{C0}	Natural frequency of the CP at the stapes [rad/s].
$\omega_C(x)$	Natural frequency of the CP at location x [rad/s].
*	Denotes the complex conjugate.
$\otimes$	Denotes convolution.

Notes:

1. complex quantities appear in bold typeface.
2. variables expressed above as functions of f may also appear as functions of η.

1. Introduction

1.1 The Anatomy and Physiology of the Human Cochlea

In this section the anatomy and physiology of the cochlea are briefly described. A detailed review of our current knowledge of the human cochlea is reported in Pickles (1988) and Dallos *et al.* (1996).

The cochlea comprises three fluid filled channels, or scalae, called the scala vestibuli, the scala media and the scala tympani. These are about 35 mm in length and are coiled into a spiral of approximately 2.5 turns about a bony core called the modiolus. The spiral structure is embedded in the temporal bone, which forms part of the outer wall of the three scalae. The scala media is separated from the scala vestibuli by Reissner's membrane, and from the scala tympani by the basilar membrane (BM). All three scalae contain fluids whose mechanical properties are similar to water and whose ionic composition is important in cochlear function. The scala vestibuli and scala tympani are connected to each other by a small port called the helicotrema, located at the apex of the spiral, whilst the scala media is closed at the apex, and is not directly connected to the other two scalae. Acoustic vibrations in the outer ear canal are transmitted by the middle-ear ossicles to a membrane at the basal end of the scala vestibuli called the oval window. The scala tympani is closed by a membrane called the round window, which faces the middle cavity.

The BM comprises radial fibres which are connected at their inner end to a bony ledge projecting from the modiolus called the osseous spiral lamina. The outer ends of the fibres connect to a structure called the spiral ligament. The stiffness of these fibres is important in determining the mechanical behaviour of the cochlea.

From a cochlear mechanical viewpoint, the spiral shape of the cochlea is unimportant and it is therefore often treated as though it were unwound into a long, straight structure. The cochlea shows the following variations in its dimensions and properties along its length. Firstly, the BM becomes broader, and mechanically more compliant from base to apex. Secondly, both the width of the osseous spiral limbus and the cross section of the scala media taper in the opposite direction to the BM, becoming narrower from base to apex. These variations form the basis of the mechanical frequency analysis that occurs in the cochlea in which stimulus frequency is mapped to location on the BM.

Acoustic stimuli in the ear canal cause vibrations of the oval window which generate a hydromechanical wave involving transverse motion of the BM. The wave, known as the travelling wave (TW) travels along the BM carrying energy from base to apex. For a pure tone stimulus, the TW amplitude response varies with position along the BM, peaking at a

point which is dependent on the stimulus frequency. The location of the peak in the response is near the apex of the cochlea for low frequency stimuli, and moves towards the base for progressively higher frequency stimuli.

The BM supports a cellular structure called the organ of Corti, which comprises support cells and two types of sensory cells: the inner and outer hair cells (IHC and OHC respectively). Above these is a gelatinous structure called the tectorial membrane, which is attached at its inner edge and which runs along the length of the cochlea. A bundle of hairs called the stereocilia projects from the top of each hair cell towards the tectorial membrane, with the tips of the OHC stereocilia being embedded in its underside. In response to vibrations of the BM, the stereocilia deflect, thereby modulating the rate of flow of ions from the surrounding cochlear fluid into the hair cells.

In the IHCs, this flow of ions into the cell initiates nerve impulses which travel along the auditory nerve to the brainstem. In this way, the vibration of a point on the BM becomes encoded as a neural signal.

The role of the OHCs is quite different from that of the IHCs: rather than detecting BM vibration, the OHCs actively amplify the vibration, thus enhancing the sensitivity of the auditory system. This amplification mechanism is not well understood, but according to the most widely accepted theories, a deflection of their stereocilia causes the OHCs to undergo length changes, thereby applying excitation forces to the BM and injecting energy into the passing TW. Since these length changes are both initiated by, and have an influence on the passing TW, a feedback loop exists involving mechanical to electrical transduction by the stereocilia, followed by electrical to mechanical transduction in the body of the OHC. The entire system is known as the cochlear amplifier, and is powered by a metabolic energy supply which maintains the ionic composition of the cochlear fluids. The action of the cochlear amplifier not only increases the sensitivity of the cochlea, but also improves frequency selectivity by sharpening the peak of envelope of the TW response. The cochlear amplifier is physiologically vulnerable, and its impairment is implicated in most cases of sensorineural hearing loss.

The mechanical response of the cochlea shows a strong compressive nonlinearity such that the ratio of amplitude of the BM vibration to the that of the acoustic stimulus reduces with increasing stimulus level. This nonlinearity greatly complicates the mechanical response since the response at any given frequency may be influenced by many frequency components in the stimulus, not just the component at the response frequency. It is thought that this nonlinearity arises predominantly from the transduction processes in the OHCs.

1.2 Otoacoustic Emissions

Otoacoustic Emissions (OAE) are low level acoustic signals that are generated in the cochlea, propagate through the middle ear and appear in the ear canal where they can be measured using a probe microphone (Kemp, 1978). The mechanism of their generation is thought to involve the cochlear amplifier, and consequently OAEs provide a clinically useful noninvasive tool for assessing the state of health of the cochlea.

OAEs are usually divided into two main classes: spontaneous OAEs and evoked OAEs. Spontaneous OAEs (SOAE) are sounds measured in the ear canal in the absence of any external stimulation. They appear as one or more narrow band spikes in the power spectrum of the ear canal sound pressure, which can be measured with a miniature microphone sealed into the ear canal.

Evoked OAEs only arise when an external sound stimulus is presented to the ear. These are usually measured using a probe in the ear canal which houses both a microphone and an earphone. Evoked OAEs are usually further divided into three subclasses: transiently evoked, stimulus frequency, and distortion product OAEs, abbreviated to TEOAE, SFOAE and DPOAE respectively (Probst *et al.*, 1991). This subdivision is based purely on the type of the evoking stimulus and is not intended to imply differences in the generation mechanisms. TEOAEs are evoked using a transient stimulus such a click, tone-burst or chirp and appear as echo-like signals after the stimulus. SFOAEs are evoked using a continuous pure tone stimulus, and appear themselves as continuous pure tones at the same frequency as the stimulus. DPOAEs are evoked by the presentation of two or more continuous pure tone stimuli and appear as continuous pure tones at intermodulation frequencies. SFOAEs and TEOAEs are sometimes referred to as 'simultaneously evoked' and 'delayed evoked' otoacoustic emissions respectively though this terminology will not be used here.

SFOAEs in humans are the main subject of this project. TEOAEs are also briefly discussed, since they are thought to be closely related to SFOAEs. DPOAEs are more complicated and will not be discussed in any detail.

1.3 Overview of the Project

The purpose of this project is to investigate the mechanisms by which SFOAEs are generated in the cochlea. The current leading theory of SFOAE generation was proposed by Shera and Zweig (1993b). According to this theory, SFOAEs arise by a process in which the TW in the cochlea is both amplified by active processes and reflected by an

array of reflection sites distributed randomly along the length of the BM. An earlier theory developed by Strube (1989) postulated that the reflection sites were distributed with approximate spatial periodicity along the BM.

One consequence of Shera and Zweig's theory is that certain features of the SFOAE frequency spectrum should vary with variations in TW response. These features are related to the group delay of the OAE signal which in turn is related to a quantity known as the periodicity of the SFOAE spectrum. Specifically, Shera and Zweig's theory predicts a relationship between the periodicity of the SFOAE and both the wavelength of the TW near its peak, and the sharpness of the TW envelope. In contrast, Strube's theory predicts that the periodicity in the SFOAE is independent of the TW shape, being determined only by the spatial periodicity in the reflection sites.

The main aim of this project is to look for experimental evidence for or against Shera and Zweig's theoretical prediction. The approach that was adopted was to induce changes in the TW response whilst looking for corresponding changes in the SFOAE frequency functions. Because of the nonlinearity of cochlear responses, manipulation of the acoustic input stimuli induces changes in the TW response. Two different manipulations were used. In the first, the level of the a pure tone stimulus is varied, leading to the phenomenon known as 'self-suppression'. In the second, two tones are presented simultaneously, leading to a nonlinear interaction known as 'two-tone suppression'.

This investigation has three main parts. First, cochlear models were developed, tested, and then used to generate theoretical predictions of the changes in SFOAEs that occur during self-suppression and two-tone suppression. Second, a signal processing method was developed for analysing SFOAE signals in terms of physical quantities in Shera and Zweig's theory of SFOAE generation. Third, experimental measurements of SFOAEs were made in human subjects, and comparisons made with theory.

In the first part, the objective was to obtain testable prediction of the SFOAE behaviour predicted both by Shera and Zweig's theory, and by Strube's theory. Several different cochlear models based on the 1-dimensional longwave formulation were assembled from elements of existing published models, and were then implemented on a PC. These elements included models of the cochlear amplifier and BM impedance inhomogeneities, which act as reflection sites for the TW. In order to test the sensitivity of the model predictions to arbitrary assumptions in the models, two different variants of the cochlear amplifier were implemented: the first due to Neely and Kim (1986) and the second due to Zweig (1991). Two different spatial distributions of BM impedance inhomogeneities were also modelled: first, the random distribution postulated in Shera and Zweig's theory (1993b); second, the periodic distribution postulated in Strube's theory

(1989). A nonlinear analysis technique (the 'quasilinear' method) due to Kanis and de Boer (1993b) was also included in the models in order to allow self-suppression and two-tone suppression to be modelled.

The major theoretical results of interest are as follows. Models based on Shera and Zweig's theory predict that the SFOAE periodicity should show specific changes during self-suppression and two-tone suppression. No such change in periodicity is predicted by models based on Strube's theory. These results provide a means of testing the two theories experimentally. Also, qualitatively similar results are obtained regardless of the different assumptions made in the model, such as the details of the cochlear amplifier characteristics. This is in agreement with Zweig and Shera's assertion (1995) that SFOAEs will arise whenever a 'tall-and-broad' TW encounters a random array of reflection sites.

In the second part, a signal processing method for accurately quantifying the SFOAE periodicity was developed. This is required because, in general, neither measured nor predicted SFOAE spectra show perfect spectral periodicity. Instead they exhibit a degree of random amplitude and phase modulation with frequency. In Zweig and Shera's cochlear model, this arises from the underlying random nature of the reflection sites on the BM. It can be shown that, as a consequence of this random element, the changes in periodicity predicted by Zweig and Shera's cochlear model are difficult to discern using previously published methods. For cochlear models, this random element can be removed by performing averaging across an ensemble of models, each with a different realization of the random reflection sites. This yields the parameters which describe the population from which the realizations were drawn. However, performing the corresponding ensemble averaging process is not possible for experimental data, where the ensemble would have to be formed from different subjects. The problem there is that, in contrast to the ensemble of cochlear models, different subjects differ in many more ways than simply their distributions of reflections sites. Therefore a method is required for estimating the periodicity of a single SFOAE spectrum. The method developed in this project is based on parametric spectral analysis. In this method, called the 4-parameter model, the measured SFOAE function is treated as if it arose from a random process with two main controlling parameters. The first of the four parameters, called here the ϕ-centre value is used to define the SFOAE periodicity in this thesis. The second parameter, called here the ϕ-bandwidth, is related to the spread of periodicities in the SFOAE signal. The 4-parameter model was tested against the results from averaging across an ensemble of cochlear models, and was found to be capable of detecting the changes in periodicity (or

ϕ-centre value) predicted by Shera and Zweig's model. Furthermore, its performance was significantly better than that of alternative estimators.

In the third part, experiments were designed and conducted to measure the behaviour of the SFOAE periodicity during self-suppression and two tone suppression. SFOAEs were recorded in 20 subjects with normal hearing under varying degrees of self-suppression and two-tone suppression. Where SFOAEs could be measured with sufficient accuracy, the 4-parameter model described earlier was used to estimate the periodicity. These results were then compared with theoretical predictions from the cochlear models.

The experimental results show that the SFOAE periodicity is altered both during self-suppression and two-tone suppression. Furthermore, the measured relationship between the periodicity and levels of the acoustic stimuli closely resembles the predictions of cochlear models based on Shera and Zweig's theory (1993b). In contrast, the experimental results contradict the predictions of models based on Strube's theory (1989). This general result holds for several different variants of cochlear model, indicating that the result is not dependent on the exact details of the models. It is concluded that the experimental results provide strong support for Shera and Zweig's theory (1993b).

1.4 Project Motivation and Scope

The two very basic questions "How does the cochlea work?" and "How are OAEs generated?" have still not been fully answered at the macromechanical level. Although plausible theories have been developed to answer these questions, these theories have yet to be fully substantiated. One motivating factor for this project comes from the desire to test these theories experimentally.

A related question is "What do OAEs tell us about the cochlea?". The answer to this question has implications not just for basic hearing science, but also for clinical audiology. Although OAEs have become a useful clinical tool in determining the state of health of the cochlea, exactly what information is contained in OAE signals remains unclear. Consequently, current clinical procedures use the most basic features of OAE signals, such as the overall level of the emission. Thus, the motivation for this project comes not just from basic hearing science, but also from clinical audiology. Though in this project no direct clinical application is sought, it is hoped that the further understanding of the cochlear mechanisms of OAE generation that is gained will lead to the development of improved clinical techniques.

In particular, this thesis is motivated by the predictions of Shera and Zweig's theory (1993b), that SFOAE periodicity is related to the shape of the TW. This prediction

differentiates the theory from an earlier theory due to Strube (1989) and thus provides a possible means of testing the two theories experimentally. Also, if the prediction is correct, it provides a link between cochlear mechanical behaviour and OAE signal characteristics.

The potential benefits of this work are in three areas. Firstly, at the purely phenomenological level, it may lead to improved methods of characterising measured OAEs, thereby revealing new patterns in OAE traces. Secondly, at the cochlear mechanical level, it tests some aspects of the leading theories of the OAE generation in humans, where there is a dearth of experimental data. Thirdly, it may point to ways of extracting useful information about a particular cochlea (such as the width of the auditory filter) from measurements of OAEs.

The scope of this investigation has been limited to SFOAEs. The reason for this is that the theoretical treatment of SFOAEs at the cochlear mechanical level is simpler than that of other OAEs. This is because the cochlea is nonlinear, and therefore frequency components in the stimulus (or in the response) can potentially interact with each other. As SFOAEs arise from a single pure tone stimulus, nonlinear effects are minimised. According to current theories, SFOAEs share their origins with other OAEs. Therefore, the investigations carried out in this project may provide insight into the generation mechanisms of other OAEs. This is important, since current clinical procedures use TEOAEs and DPOAEs.

1.5 Overall Aims and Objectives

1.5.1 Aims

The general aim of this project is to further the understanding of cochlear mechanics by exploring the link between measurable characteristics of OAEs and current macromechanical theories of OAE generation. In particular, the aim is to answer the following questions:

1. What predictions do the theories of Shera and Zweig (1993b) and of Strube (1989) make about the variation of SFOAE periodicity during self-suppression and two-tone suppression, and how can these variations be qualified?
2. How does the SFOAE periodicity measured in human ears vary during self-suppression and two-tone suppression?
3. Do the experimental data concur with the predictions of the theories of either Shera and Zweig or of Strube?

1.5.2 Deliverables

As part of the process of answering these questions, the aim is to deliver:

1. cochlear models which capture the essential features of the two theories, and which are capable of predicting SFOAEs during self-suppression and two-tone suppression,
2. appropriate predictions from the cochlear models,
3. a signal processing method for quantifying SFOAE periodicity,
4. experimental measurements of SFOAEs during self-suppression and two-tone suppression.

1.5.3 Objectives

The project can be broken down into the following objectives:

1. To develop and implement macromechanical cochlear models in Matlab, which capture the essential features of both Shera and Zweig's theory and of Strube's theory. These models must be capable of the showing the effects of self-suppression and two-tone suppression both on the TW envelope and on SFOAEs.
2. To examine the performance of various versions of the models in detail, in order to assess the sensitivity of the results to various modelling assumptions. The effects of differences in the formulation of the middle ear model, in the properties of the BM, and in the characteristics of the cochlear amplifier are all investigated.
3. To use the cochlear models to make predictions of the effect on SFOAE spectra of self-suppression and of two-tone suppression.
4. To develop a signal processing method for quantifying the periodicity of SFOAE spectra. Problems arise in quantifying SFOAE periodicity because of the complexity of the SFOAE spectra predicted by Shera and Zweig's model. The signal processing method is to be designed to extract meaningful parameters from the SFOAE spectra predicted by Shera and Zweig's model, in order to allow comparisons to be made with measured data.
5. To design and carry out experiments to measure the effect on SFOAEs in humans of self-suppression and of two-tone suppression.
6. To use the signal processing method to quantify both the measured and theoretical variations in SFOAE periodicity during self-suppression and two-tone suppression.

7. To refine the cochlear models and signal processing methods further where necessary, based on the experimental results.

1.6 Structure of the Thesis

This thesis is broken into four parts. In the first part, the published literature is reviewed on the measurement of SFOAEs and on the theories of SFOAE generation. In the second part, the development of a cochlear model, and of a signal analysis method are described, leading to the generation of testable theoretical predictions. In the third part the experiments are described and comparisons are made between the experimental results and the theoretical predictions. Finally, in the fourth part the overall implications of the results are discussed.

PART I: LITERATURE REVIEW

2 Experimental Findings on OAEs

2.1 Definition of OAEs

OAEs are defined by Probst *et al.* (1991) as 'acoustic energy produced by the cochlea and recorded in the outer ear ...'. In the case of SOAEs, this definition is unproblematic. However, in the case of evoked OAEs the application of this definition is not straightforward. This is because the pressure in the ear canal comprises several components: the direct wave from the external stimulus, reflections originating from the eardrum or middle ear, plus the desired OAE originating in the cochlea. This is a particular problem for TEOAEs and SFOAEs, where the OAE and stimulus signals overlap in both the time and the frequency domains. One way of defining an evoked OAE (at least conceptually) is as the additional acoustic signal in the ear canal that arises as a results of the active processes in the cochlea. Practical methods of measuring TEOAEs and SFOAEs are discussed later.

SFOAEs are the main subject of this project. Recall that SFOAEs are continuous pure tones evoked using a continuous pure tone stimulus at the same frequency. Due to cochlear nonlinearity, a single pure tone stimulus will evoke not only a pure tone response at the stimulus frequency, but also pure tone responses at other frequencies, such as integer multiples of the stimulus frequency. These components are not classed as SFOAEs (and are in any case generally quite small).

2.2 A Review of the Theory and Analysis of Reflections

Before reviewing the experimental data on OAEs, it is useful to consider some very simple reflection phenomena. This anticipates the "reflection hypothesis" which was introduced by Kemp (1978) to explain the generation of OAEs, and which is reviewed in more detail later. In this section, some of the basic physics of reflections, together with some useful signal processing techniques are introduced.

Fig. 2.1, panel (a) shows a simple arrangement for generating a simple reflection in a duct. A loudspeaker generates a single click, which travels as a one-dimensional wave along the duct, is reflected at the right hand wall, travels back down the duct and is completely absorbed at the left hand wall.

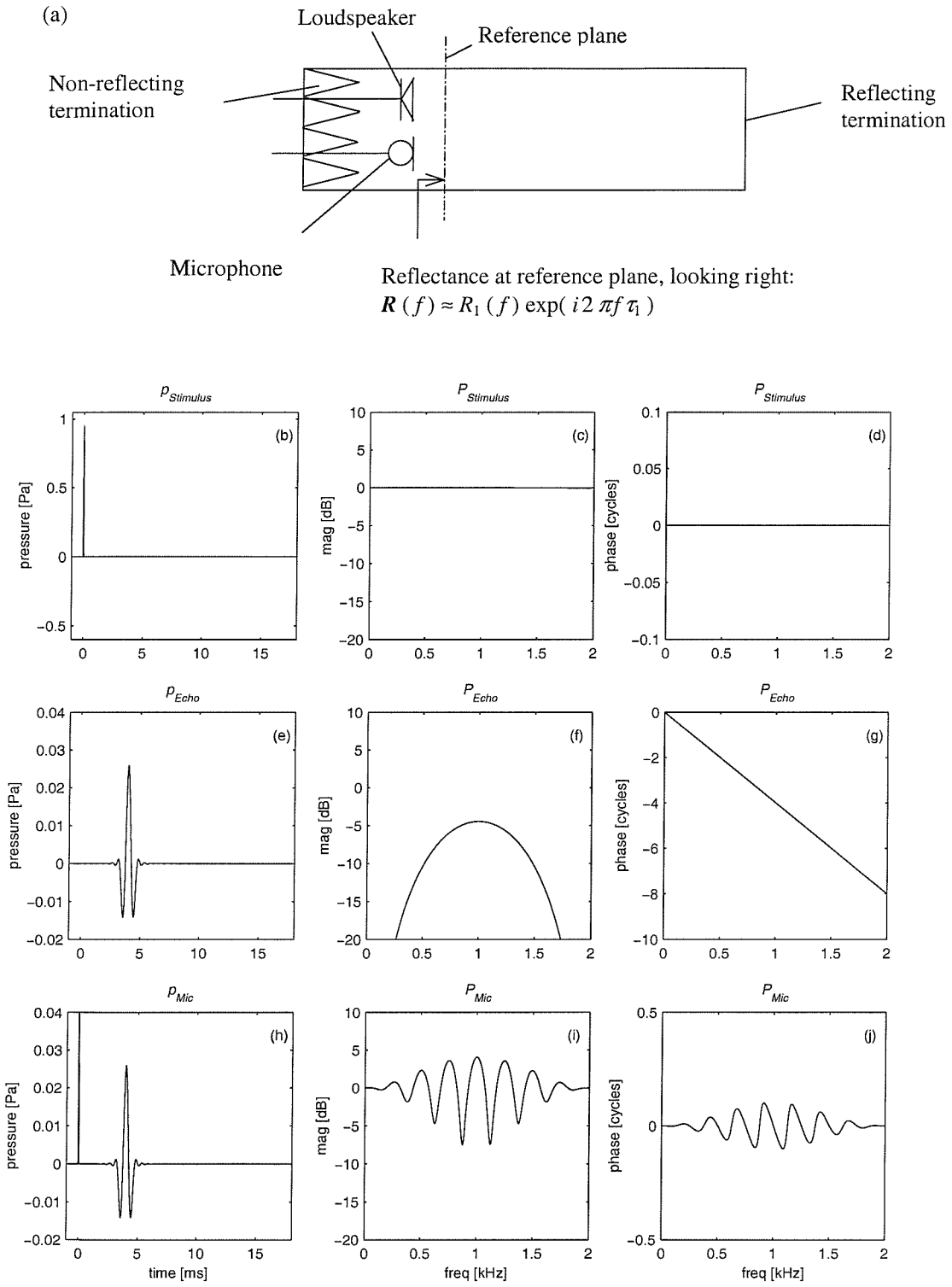


Fig 2.1. A single reflection in an acoustic cavity. Time and frequency domain representations of the microphone response arising from a single reflection. The microphone, loudspeaker and reference plane in panel (a) are all coincident with the left hand wall of the duct, but are drawn separated for clarity. The microphone signal, panel (h), can be represented as the sum of two components: the "Stimulus" component, panel (b), and the "Echo" component, panel (e). The stimulus component may be defined as the microphone response which would have been measured had the right-hand duct termination been completely non-reflecting. In this example, the magnitude of the reflectance, $R_1(f)$, is broad band, peaking at 0.6, and the time delay, τ_1, is 4 ms. Middle panels (c), (f) and (i) show the magnitude of the Fourier transforms of the left hand panels (b), (e) and (h) respectively. The right hand panels, (d), (g) and (j) show the phase of the Fourier transforms of the left hand panels (b), (e) and (h), respectively. The echo and stimulus are related in the frequency domain via the (complex) reflectance: $\boldsymbol{P}_{Echo}(f) = \boldsymbol{P}_{Stimulus}(f)\,\boldsymbol{R}(f)$.

Here it is assumed that the combination of the loudspeaker, microphone and left-hand termination is non-reflecting, that the system is linear, that the side-walls are rigid and that only plane waves propagate in the duct. Also, the right hand wall is assumed to reflect all incident waves in a broad frequency band between about 0.25 and 1.75 kHz, without introducing any significant delay[1], whilst the passage of the wave along the duct introduces delay, but with no significant change in amplitude. The click passes the microphone twice before being absorbed, giving rise to two pulses in the microphone pressure signal, panel (h). This signal can be split into two components: the stimulus component and the echo component. The two components of the microphone signal are shown in panels (b) and (e), and their Fourier transforms are shown to the right of these. The stimulus is a very short click at zero time, whose Fourier transform therefore has a constant magnitude and zero phase over a wide frequency band. The Fourier transform of the echo component has a broad band magnitude arising from the reflection characteristics of the right hand wall, and a linear phase arising from the transmission delay.

To a good approximation, the sound field in the duct satisfies the simple one-dimensional wave equation. From this equation, it can be shown that any arbitrary sound field in the duct can be represented as the superposition of a right and a left going wave, referred to here as the forward and backward waves respectively. Consequently, it is useful to define a quantity known as the reflectance (or reflection coefficient), which, at any frequency, is defined as the ratio of the complex amplitude of the reflected wave to that of the incident wave. In panel (a) the reflectance at the plane of the microphone looking in the forward direction is given by:

$$\text{Reflectance, } \boldsymbol{R}(f) \equiv \boldsymbol{P}_{Backward}(f) / \boldsymbol{P}_{Forward}(f) \qquad [2.1]$$

where symbols in bold typeface are used to denote complex quantities, and where $\boldsymbol{P}_{Backward}$ and $\boldsymbol{P}_{Forward}$ are the complex pressures at the microphone due to the backward and forward components. For this simple case, the forward going wave is simply the stimulus wave, whilst the backward going wave is simply the echo component. This leads to the following simple relationship:

$$\boldsymbol{P}_{Mic} = \boldsymbol{P}_{Stimulus} + \boldsymbol{P}_{Echo} = \boldsymbol{P}_{Stimulus}(1+\boldsymbol{R}) \qquad [2.2]$$

where all quantities are assumed to be complex functions of frequency.

Because the Fourier transform of the stimulus is uniform (i.e., independent of frequency) (panels c and d), the reflectance is identical in shape to the Fourier transform of the echo, shown in panels (f) and (g). The magnitude of the reflectance thus is related to

[1] In order for this broad band filtering to be physically realizable, there must be some non-zero delay, but this can be very small compared to the transmission delay.

the broad band filtering that occurs between the incident and the reflected waves, whilst the phase is related to the delay. The linearity of the phase indicates that there is no dispersion: all frequency components are delayed by the same time. The slope of the phase indicates the actual time delay:

$$\tau_{GD}(f) \equiv -\frac{1}{2\pi}\frac{\mathrm{d}\theta(f)}{\mathrm{d}f} = \tau_1 \qquad [2.3]$$

where τ_{GD} is the delay in seconds (known formally as the group delay), θ is the phase angle in radians, and f is the frequency in Hz. In this case τ_{GD} is independent of frequency and is denoted τ_1.

Though it is conceptually useful, the reflectance is not directly measurable. Instead, measurements yield only $\boldsymbol{P}_{Mic}$ for a given loudspeaker input voltage. However, given knowledge of the loudspeaker source impedance[2], $\boldsymbol{P}_{Stimulus}$ could be calculated from this input voltage, and hence the reflectance can be solved from [2.2] (Keefe, 1997).

In addition to the reflectance, it is useful to examine the transfer function given by $\boldsymbol{P}_{Mic} / \boldsymbol{P}_{Stimulus}$. The form of this function resembles the frequency response shown in panels (i) and (j). Here, the echo causes a ripple pattern to be superposed onto the magnitude spectrum of the microphone pressure (panel i). The ripple spacing is related to the delay between the stimulus and the echo, by $\tau_1 = 1 / \Delta f$, where Δf is the frequency interval in Hz between adjacent ripple peaks. Physically, these ripples arise from the interference between the stimulus and echo at the location of the microphone. This can be understood with reference to fig. 2.2, which shows the complex phasor representations of the stimulus and echo components for a single frequency. (See Kemp and Chum, 1980a, fig 1. and Randall, 1987, fig 8.2) The phasor representing the echo pressure lies at an angle of $2\pi f \tau_1$ relative to the real axis, and therefore rotates anticlockwise with increasing frequency. The length of the phasor changes only slowly with frequency. This causes a roughly periodic change in the length and angle of the microphone phasor, corresponding to ripples in the magnitude and phase spectra of the microphone signal, shown in panels (i) and (j).

A more complicated situation arises when the left hand wall is reflecting, as is shown in fig. 2.3, panel (a). Again the loudspeaker, microphone, reference plane and left hand wall are coincident. Two reflectances are now defined at the reference plane: $\boldsymbol{R}(f)$ for incident waves approaching in the forward direction, and $\boldsymbol{r}(f)$ for those approaching in

[2] It is assumed that the loudspeaker acts as a volume velocity source (proportional to voltage) shunted by an acoustic impedance. To achieve a reflectionless termination, this acoustic impedance must equal the characteristic impedance of the medium (scaled appropriately with the cross sectional area of the duct).

the backward direction. This leads to a pressure pulse travelling back and forth, being reflected at each encounter with the termination (e.g., Keefe, 1997).

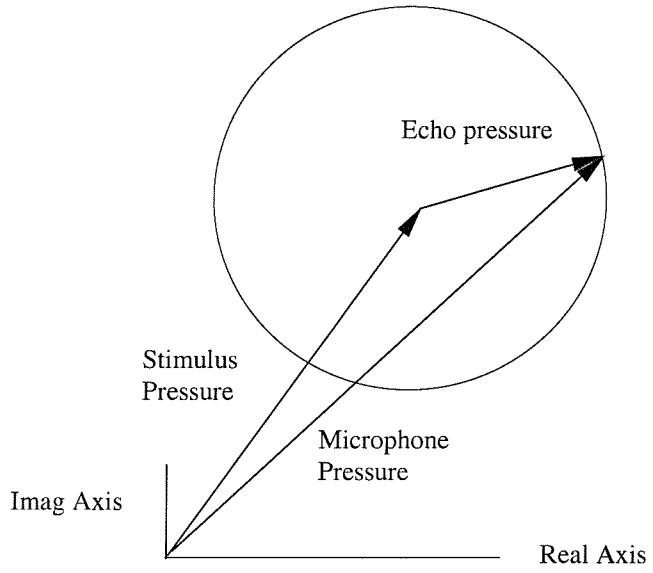


Fig 2.2. Phasor representation of the stimulus and echo components for a simple reflection, based on Kemp and Chum, 1980a, fig 1, by permission of D. T. Kemp and Delft University Press.

As before, it is convenient to split the microphone signal in panel (h) into the stimulus and echo components, shown in panels (b) and (e), where the stimulus component is defined as the pressure signal that would have been measured had the right hand wall been non-reflecting. The echo component is simply defined as the microphone signal minus the stimulus. Now the forward and reverse pressure components no longer correspond to the stimulus and echo components, because the echo component is made up of forward as well as reverse waves (the stimulus component comprises only the first forward wave). This means that the echo spectrum in panels (f) and (g) no longer resembles the reflectance, $\boldsymbol{R}(f)$. It can be shown[3] that the following relationship holds

[3] This can be verified by noting that the ratio of the forward to the reverse echo component in [2.4] equals $\boldsymbol{r}(f)$, whilst the ratio of the reverse echo component to the sum of forward echo component and the stimulus component equals $\boldsymbol{R}(f)$.

between the echo and stimulus components:

$$P_{Echo} = \frac{P_{Stimulus} R(1+r)}{1-rR} = P_{Stimulus} R(1+rR+(rR)^2+\ldots \text{ (reverse)} \quad P_{Stimulus} rR(1+rR+(rR)^2+\ldots \text{ (forward)} \qquad [2.4]$$

where explicit dependence on frequency has been omitted. The binomial expansion of the denominator emphasises the interpretation of the echo signal as a sum of multiple reflections. The value of $r(f)$ depends only on the source impedance. When $r = 0$, [2.4] simplifies to the single reflection case described earlier. For simplicity, $r(f)$ has been assumed real and constant in the example shown in fig 2.3.

Unlike the single reflection case, the spectrum of the echo component in panels (f) and (g) now shows a ripple pattern, which arises from the periodicity in the time domain signal in panel (e). In fact the ripple peaks correspond to the acoustic natural frequencies of the system, which occur as a series of harmonics. The fundamental frequency, and the spacing between the higher harmonics are given by the reciprocal of the time interval between successive reflections. The width of these peaks is inversely related to the decay time of the envelope of the pulse train seen in panel (e), which is determined by the damping in the system. A large value of the product $|rR|$ gives rise to a slow decay of the pulse train, and therefore sharper ripples in the frequency domain. Note also that the simple phasor diagram (fig. 2.2) no longer strictly applies, since the echo component is now made up of an infinite sum of phasors, each one rotating (as frequency increases) at a different rate, depending on its delay.

The ripple patterns in the spectra of the echo (panel (f)) and microphone (panel (i)), resemble each other because the time delay between multiple reflections is the same as that between the stimulus and the first reflection. This arises because the loudspeaker and the left hand wall are coincident, and because $r(f)$ has zero phase.

Another quantity of interest is the acoustic impedance looking forwards at the plane of the microphone, which is defined as the complex ratio of the acoustic pressure to the volume velocity. A standard result relating the impedance, $Z(f)$, to the reflectance is:

$$Z = \frac{Z_0(1+R)}{1-R} \qquad [2.5]$$

where Z_0 is the characteristic impedance of the medium. $Z(f)$ is independent of the reflectance, $r(f)$, since it depends only on the system to the right of the reference plane.

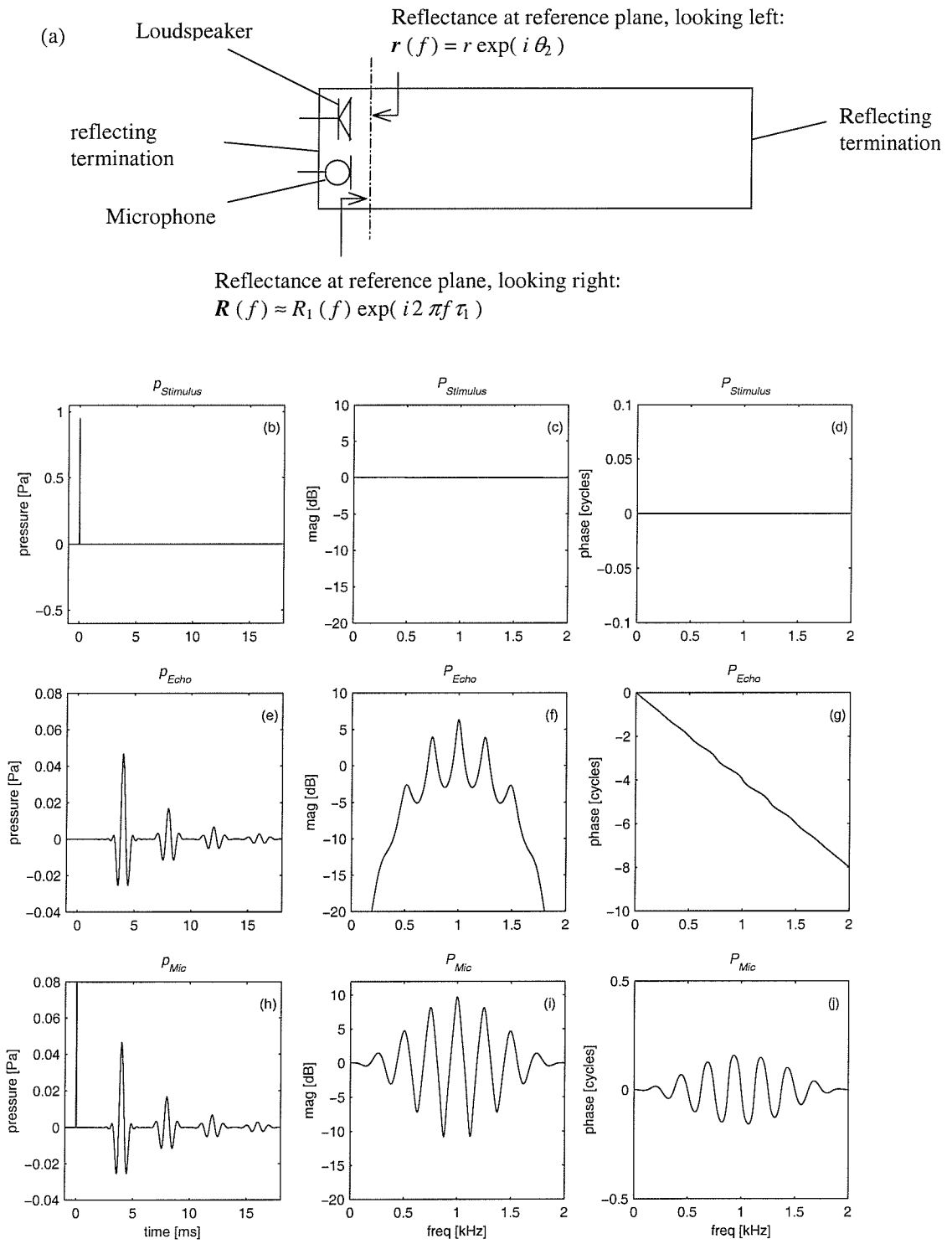


Fig 2.3. Multiple Reflections in an acoustic cavity. Panels (a)-(j) are as in fig 2.1, but with a reflecting left hand wall. The "stimulus" component in panel (b) is defined as the microphone response that would have been measured had the right-hand duct termination been completely non-reflecting. For the reflectance looking right, the magnitude, $R(f)$ is broad band with a peak value of 0.6, and the time delay, $\tau_1 = 4$ ms. For the reflectance looking left, the magnitude, $r(f) = 0.8$, and the phase angle, $\theta(f) = 0$ for all f. Middle panels (c), (f) and (i) show the magnitude of the Fourier transforms of the left hand panels (b), (e) and (h) respectively. The right hand panels, (d), (g) and (j) show the phase of the Fourier transforms of the left hand panels (b), (e) and (h), respectively. The echo and stimulus are related in the frequency domain via the two (complex) reflectances:

$\boldsymbol{P}_{Echo}(f) = \boldsymbol{P}_{Stimulus}(f)\,\boldsymbol{R}(f)\,[\,1 + \boldsymbol{r}(f)\,] / [\,1 - \boldsymbol{r}(f)\,\boldsymbol{R}(f)]$.

Now the transfer function between the microphone output and the loudspeaker input will approximate this acoustic impedance, provided that $\boldsymbol{r}(f) \approx 1$, as is the case here. This can be seen by setting $\boldsymbol{r}(f) = 1$ in [2.4], and by noting that, for a high impedance loudspeaker[4], the volume velocity is proportional to the voltage input. This means that the spectrum shown in panels (i) and (j) closely resembles the acoustic impedance seen at the microphone. Therefore the peaks in the ripple pattern correspond closely to antiresonances of the system, defined as frequencies where the input resistance is large.

Summarising the main points from these two simple cases:

1. For the single reflection case, the transfer function between echo and stimulus signals equals the reflectance at the transducers, looking forwards. The slope of the phase of the reflectance is determined by the time delay between stimulus and echo.
2. For the single reflection case, the transfer function between microphone and loudspeaker shows a ripple pattern whose spacing is determined by the time delay between stimulus and echo.
3. For the multiple reflection case, the transfer function between echo and stimulus is no longer simply the reflectance looking forwards. Ripples appear in both its the magnitude and phase. The average slope of the phase of the reflectance is still determined by the time delay between stimulus and echo.
4. For the multiple reflection case, the transfer function between microphone and loudspeaker closely resembles the acoustic impedance at the microphone, provided that the source impedance is high, such that the reflectance looking backwards has a value close to 1.
5. To recover the reflectance from measurements of $\boldsymbol{P}_{Mic}$ a knowledge of the loudspeaker source impedance is required.

It should be noted that in both cases, the presence of ripples in the $\boldsymbol{P}_{Mic}$ spectrum does not imply that the reflectance at the right hand wall is higher at the frequencies of the ripple peaks. Instead, the ripple peaks indicate constructive interference between two or more temporally separated components in the signal.

2.3 A Qualitative Description of Measured TEOAE and SFOAE Signals

TEOAEs were first measured by Kemp (1978). He presented a click to the ear canal via an earphone, and then measured the resulting acoustic pressure in the ear canal using a

[4] The impedance must be high for $\boldsymbol{r}$ to approach 1. A similar equation to [2.5] can be written relating $\boldsymbol{r}$ to the source impedance.

miniature microphone. In fact, in order to improve the signal-to-noise ratio, a train of clicks was presented, and then synchronous averaging performed. The TEOAE appeared in the microphone signal as an echo-like component typically lasting around 20 ms. Kemp (1978) proposed that TEOAEs originate in the cochlea, and that they involve active cochlear processes. This has since been verified by a large body of evidence reviewed by Probst *et al.* (1991) and discussed later.

The TEOAE signal has several characteristics that are different from those of the simple echoes discussed in section 2.2. Of particular importance here are 'frequency dispersion' and 'compressive nonlinearity'. Frequency dispersion means that the different frequency components in the TEOAE have different delays, such that higher frequency components appear before lower frequency components (Kemp, 1978). This means that, unlike the simple echoes in section 2.2, the shape of the TEOAE waveform is quite different from that of the stimulus. Compressive nonlinearity means that the slope of the input-output function, defined by plotting the rms amplitude of the TEOAE waveform against that of the evoking click, is less than 1 dB/dB (Kemp, 1978). Again this is unlike the simple behaviour of acoustic reflections in a duct (though this cannot be seen simply by examining the single waveforms in section 2.2). Both these phenomena are discussed further in part II in the context of cochlear macromechanical models.

Shortly after Kemp's discovery of click evoked OAEs, it was reported that pure tone stimuli also evoked emissions (Kemp and Chum, 1980a; Wilson, 1980a). These are referred to here as SFOAEs. They are seen most readily in a normally hearing subject by exciting the ear canal with a continuous low-level pure tone stimulus using an earphone with a high impedance, and then measuring the resulting ear canal pressure. As the stimulus frequency is varied, a ripple is observed in the amplitude of the frequency spectrum of the ear canal pressure. Kemp and Chum (1980a) attributed these ripples to alternate constructive and destructive interference between the (pure tone) stimulus wave and the delayed OAE wave. This is the same interference mechanism that gave rise to the ripple patterns seen in the spectra shown in figs. 2.1(i) and 2.3(i), and represented by the phasor diagram in fig. 2.2. As will be discussed later, SFOAEs also exhibit frequency dispersion and compressive nonlinearity.

If the generating mechanism for OAEs were entirely linear then there would be no theoretical need to define SFOAEs independently of TEOAEs. This is because the Fourier transform of the click evoked OAE would be identical to the pure tone evoked OAE. Since both TEOAEs and SFOAEs show nonlinearity, as seen in their input-output functions, SFOAEs cannot be predicted from TEOAEs (or vice versa). Nevertheless, the

two classes of OAE do share many properties, including frequency dispersion and compressive nonlinearity (Kemp and Chum, 1980a), which will be discussed further.

2.4 The Evidence for Active Processes

In section 2.1, an OAE was defined as the component in the signal arising from active cochlear processes. It has also been stated the ripples seen in the spectrum of the ear canal impedance indicate the presence of echo-like components. However, whether or not these components arise from active processes, and therefore qualify as OAEs, cannot be established from observations of the signals alone. In fact, there is now extensive experimental evidence, reviewed by Probst *et al.* (1991), to show that the ripples arise from active processes. Most notable is the fact that no ripples are measured in ears with known cochlear hearing loss exceeding about 30 dB HL (Probst *et al.*, 1991). Another important experimental observation is the that the ripple exhibits a compressive nonlinearity such that the amplitude of the ripple pattern reduces as the stimulus level increases, until at around 60 dB SPL it is barely noticeable (Kemp and Chum, 1980a). Such compressive nonlinearity is also seen in other types of OAEs in humans (Probst *et al.*, 1991) and in *in vivo* measurements of cochlear mechanical responses in animals (e.g., Sellick *et al.*, 1982). It is generally believed that the passive cochlear mechanics are virtually linear and that the compressive nonlinearity arises from saturation of the cochlear amplifier (Dallos, 1996).

In addition to this experimental evidence, theoretical considerations suggest that the passive cochlea will be reflectionless[5], and would not give rise to the observed ripples (de Boer, 1996; Shera and Zweig, 1991a, 1991b, 1993a).

2.5 Frequency Dispersion of SFOAEs

As for the simple case of an echo seen in fig. 2.1(i), the ripple spacing in the magnitude of the ear canal pressure frequency spectrum depends on the time delay of the SFOAE component with respect to the stimulus component. Measurements show that the peaks in the ripples become more widely spaced as the stimulus frequency is increased

[5] Note that the term "reflectionless" here means that the reflectance measured at the base of the cochlea looking apically is zero. This reflectance is defined in terms of forward and backward cochlear TWs, and should not be confused with the reflectance of the eardrum. The cochlea, whether passive or active, has an impedance which loads the eardrum via the middle cochlea, leading to an impedance mismatch between the air in the ear canal and the eardrum. This means that reflections will occur even for a "reflectionless" (i.e., passive) cochlea. This is discussed further in section 2.9, and in part II of this thesis.

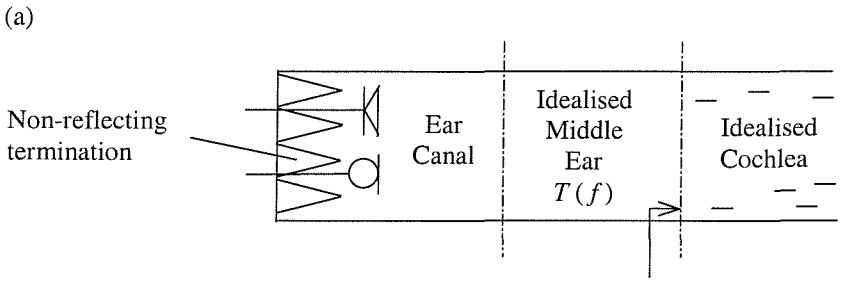
(Kemp and Chum, 1980a; Zwicker and Schloth, 1984; Dallmayr, 1987), indicating that higher frequencies have shorter delays than lower frequencies (cf., frequency dispersion in TEOAEs above). In fact, to a first approximation, it has been found that the ripple spacing is roughly proportional to the stimulus frequency (e.g., Wilson, 1980a; Zwicker and Schloth, 1984; Kemp, 1986; Dallmayr, 1987; Zweig and Shera, 1995). This holds over a frequency range of 0.5 to at least 4 kHz (Zweig and Shera, 1995). Kemp and Chum (1980a) used SFOAE phase measurements to derive the group delay, defined in [2.3], which gives a direct measure of the delay of each frequency component in the signal. This quantity also revealed dispersion similar to that indicated by the ripple spacing.

2.6 Analysis of an Idealised OAE Signal

It is useful at this point to consider a highly idealised phenomenological model of the middle ear and cochlea which has been contrived to generate SFOAEs with the property that the ripple spacing is proportional to the stimulus frequency. This idealisation, shown in fig. 2.4, is similar to Strube's analysis (1989) of a system with a group delay inversely proportional to frequency.

In fig. 2.4 the idealised cochlea is characterised by a reflectance at its base, looking forwards (i.e., apically). This has a constant magnitude given by R_1 , and has a phase angle in radians given by $\theta_1(f) = -A \log_e(f/f_{Ref})$ where A and f_{Ref} are constants. The idealised middle ear is assumed to be reflectionless and to pass all frequencies in a broad band from 0.5 to 1.5 kHz. The middle ear transmission coefficient (i.e., the ratio of transmitted wave pressure to incident wave pressure) is given by $T(f)$, and for simplicity has been assumed wholly real and identical in both directions. The ear canal is assumed to be very short, such that delays are negligible compared to delays arising in the cochlea. The ear canal termination is assumed to be reflectionless. In order to illustrate the correspondence between the ripple spacing in the frequency domain and delays in the time domain, the system is assumed to be linear. The SFOAE is then related to the stimulus by:
$\boldsymbol{P}_{OAE}(f) = \boldsymbol{P}_{Stimulus}(f)\, T(f)^2\, \boldsymbol{R}(f)$.

It must be stressed that this model is designed only to illustrate how the basic SFOAE ripple pattern is related to two properties of the system: a phase angle with a logarithmic variation, and a roughly constant magnitude over a broad band frequency band. It should also be noted that use of reflectance does not imply that the mechanism for OAE generation is passive reflection. It merely characterises the relationship between acoustic input and output waveforms (at a point in space) at a phenomenological level. The reflectance (also called the reflection coefficient) has been used in a similar way by several



Idealised reflectance at oval window, looking apically:

$\boldsymbol{R}(f) = R_1(f) \exp[\, i\,\theta_1(f)\,]$

$\theta_1(f) = -A \log_e(f/f_{ref})$

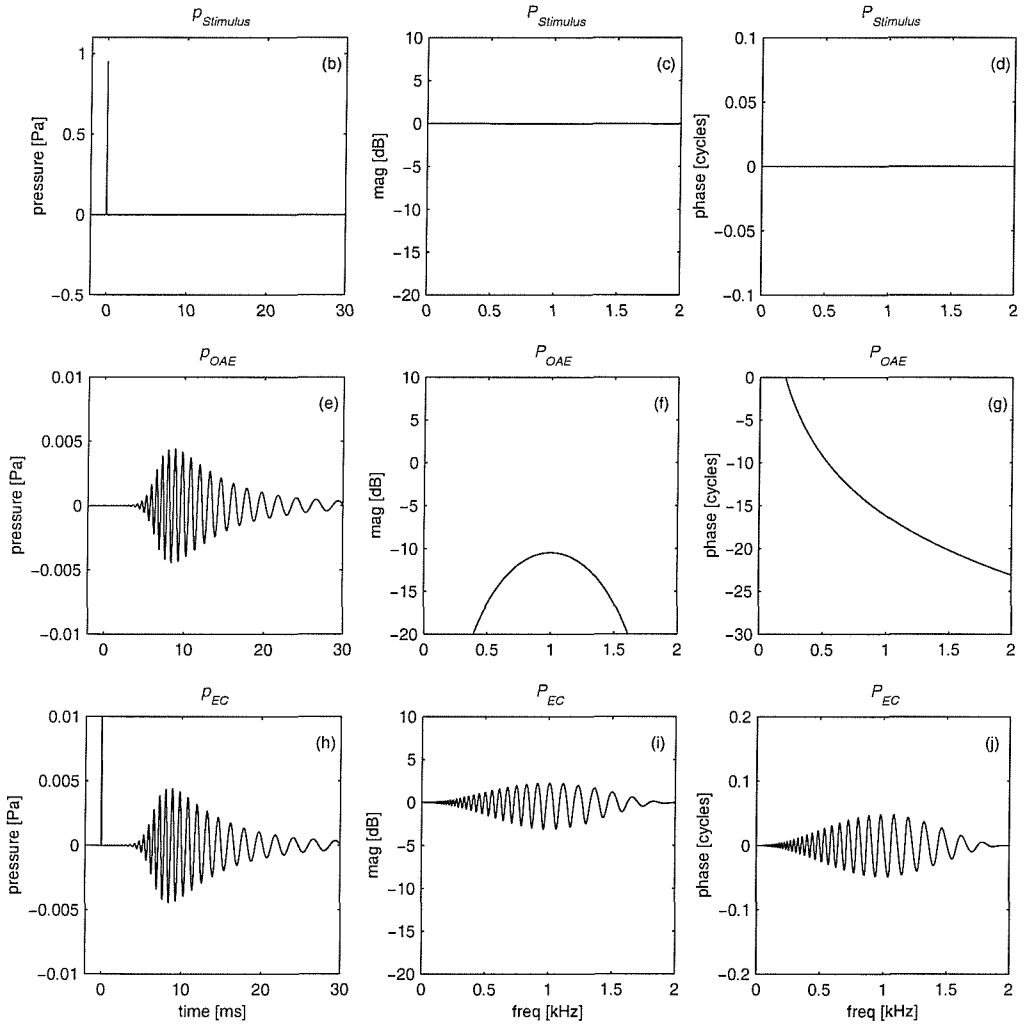


Fig 2.4. OAEs from an idealised ear, in the absence of multiple reflections. Panel (a) shows a highly idealised ear canal, middle ear and cochlea. The ear canal is closed at its left-hand end with a non-reflecting termination. The idealised middle ear is perfectly non-reflecting, though not perfectly transparent. It has a broad band transmission coefficient characterised by the attenuation function, $T(f)$. This applies to both left and right travelling signals, giving rise to the magnitude spectrum shown in panel (f). Phase changes introduced by the middle ear are neglected. The idealised cochlea has a highly simplified reflectance, $\boldsymbol{R}(f)$, whose magnitude is constant and whose phase varies logarithmically with frequency. The parameters in panel (a) are $R_1(f) = 0.3$ and $A = 2\pi\times10$ (giving a periodicity of approximately 10%). The OAE and stimulus are related in the frequency domain by: $\boldsymbol{P}_{Echo}(f) = \boldsymbol{P}_{Stimulus}(f)\, T(f)^2\, \boldsymbol{R}(f)$. The ear canal pressure is the sum of the stimulus and echo components. The stimulus component, panel (b), is defined as the microphone response which would have been measured were the cochlea completely non-reflecting.

authors (e.g., Kemp, 1980; Strube, 1989; Shera and Zweig, 1993a; Zweig and Shera, 1995). A further point to note is that the reflectance only gives a complete characterisation of the input-output relationship for a linear system. However, it may be usefully extended to give a partial characterisation of nonlinear systems (see section 4.11). More realistic phenomenological and physiological models are discussed in part II of this thesis.

For the idealised ear in fig. 2.4 panel (a), the SFOAE magnitude spectrum (panel f) shows a broad band characteristic which arises from the middle ear transmission coefficient, whilst the SFOAE phase spectrum (panel g) shows the logarithmic variation which arises from the cochlear reflectance. The ear canal magnitude and phase spectra (panels i and j) show the ripple pattern with a ripple spacing proportional to frequency.

In the time domain, panel (e) (which corresponds to an idealised TEOAE), different frequency components have different delays. As before, the delay of a particular frequency component can be quantified by the group delay, $\tau_{GD}(f)$, defined as the slope of the phase spectrum:

$$\tau_{GD}(f) \equiv -\frac{1}{2\pi}\frac{\mathrm{d}\theta_1}{\mathrm{d}f} = \frac{A}{2\pi f} \qquad [2.6]$$

where τ_{GD} is in seconds, θ_1 is in radians, and f is the stimulus frequency in hertz. Physically, the group delay of a system gives a measure of the delay between energy being delivered to the system and energy appearing at the output. This is best understood by considering the input as a tone burst signal described by a sinusoidal carrier signal of frequency, f, multiplied by a pulse shaped envelope. The output of the system is then also a tone burst signal of carrier frequency, f, but whose envelope is delayed by $\tau_{GD}(f)$ with respect to the input envelope[6].

Equation [2.6] shows that the group delay of the idealised SFOAE is inversely proportional to the frequency. The effect of this in the time domain is seen in panel (e), where high frequency components appear before low frequency components. The peak in the envelope of the waveform in panel (e) appears at a time determined by the delay of the dominant frequency component (in this case those around 1 kHz, as determined by the middle ear transmission). The onset time of the envelope is determined by the highest frequencies present (those around 1.5 kHz), since these have the shortest delay.

In the idealisation in fig. 2.4, where the middle ear and ear canal termination are reflectionless, the phase of the SFOAE relates directly to the phase of the reflectance. In reality, however, both the ear canal termination and the middle ear will be significantly reflecting. A complete analysis of this more complex case is given by Shera and Zweig

[6] In general, the shape of the output envelope differs from that of the input.

(1993a) and reviewed in part II. The intermediate case in which the ear canal termination is reflecting, but the middle ear is not, is shown in fig. 2.5. The purpose of this is to illustrate how multiple reflections manifest themselves in some of the signal representations that have been used to characterise SFOAEs. Because of dispersion, the presence of multiple discrete reflection components in the OAE signal is not immediately obvious in the time domain (panel e). However, compared to the single reflection case, fig. 2.4(e), it can be seen that the OAE waveform is distorted, and decays more slowly. Also, multiple reflections introduce ripples into the OAE magnitude and phase spectra (fig. 2.5 f and g). Apart from the middle ear transmission coefficient, the relationship between the idealised OAE and the stimulus given in equation [2.7] is identical to that for the multiple reflection case in equation [2.4].

$$P_{OAE} = \frac{P_{Stimulus} T^2 R(1+r)}{1-rT^2R} = P_{Stimulus} T^2 R(1+rT^2R+(rT^2R)^2+\ldots \text{ (reverse)} \qquad [2.7]$$
$$P_{Stimulus} rT^2 R(1+rT^2R+(rT^2R)^2+\ldots \text{ (forward)}$$

In summary, for both of these two idealised models of the ear, a ripple pattern arises in the magnitude spectrum of the ear canal pressure signal. This arises from the delay between the OAE and the evoking stimulus which is determined by the phase spectrum of the idealised cochlear reflectance. There is no ripple pattern in the magnitude of the reflectance, which has here been assumed to be independent of frequency. In the following sections, the similarities between these idealised results and actual experimental data will be discussed.

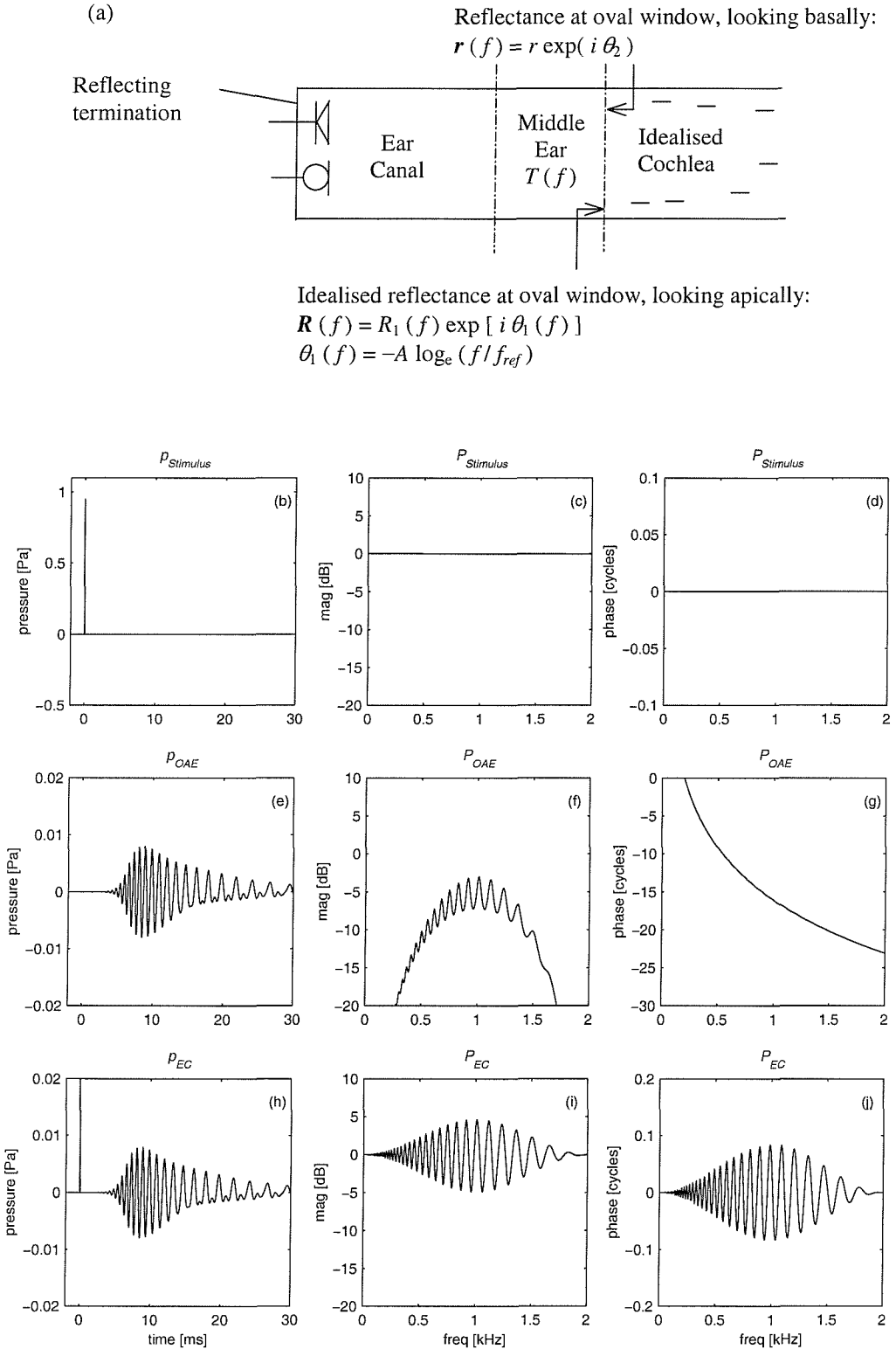


Fig 2.5. OAEs in an idealised ear including multiple reflections. Panel (a) shows a highly idealised ear canal, middle ear and cochlea. The ear canal is closed at its left-hand end with a reflecting termination. The idealised middle ear is perfectly non-reflecting, and has a broad band transmission coefficient, characterised by the attenuation function, $T(f)$. This applies to both left and right travelling signals, giving rise to the magnitude spectrum shown in panel (f). Phase changes introduced by the middle ear are neglected. The idealised cochlea has a highly simplified reflectance, $\boldsymbol{R}(f)$, whose magnitude is constant, and whose phase varies logarithmically with frequency. The parameters in panel (a) are $R_1(f) = 0.3$ and $A = 2\pi \times 10$ (giving a periodicity of approximately 10%). The OAE and stimulus are related in the frequency domain by: $\boldsymbol{P}_{OAE}(f) = \boldsymbol{P}_{Stimulus}(f)\, T(f)^2\, \boldsymbol{R}(f)\, [1 + \boldsymbol{r}(f)] / [1 - \boldsymbol{r}(f)\, T(f)^2 \boldsymbol{R}(f)]$.

2.7 Experimental Measurements of the Cochlear Reflectance

In the previous section, an idealised phenomenological model of OAEs was described which showed how the periodic ripple pattern that can be seen in the ear canal pressure might be explained by the form of the phase of the cochlear reflectance. This explanation was first put forward by Kemp (1978, 1980). However, it is conceivable that the ripples could arise from a periodic variation in the strength of the cochlear amplifier with position along the BM (e.g., Manley, 1983). In contrast to the idealised results presented above, this would manifest itself as spectral periodicity in the amplitude of the cochlear reflectance. Shera and Zweig (1993a) investigated the form of the amplitude and phase of the cochlear reflectance using SFOAEs. They made detailed measurements of the SFOAE signal characteristics and compared these with predictions from a model that included the characteristics of the OAE probe, the ear canal, the middle ear and the cochlear reflectance. Their results confirmed Kemp's hypothesis that the ripple pattern arises from the variation in the phase of the cochlear reflectance. They also showed that the magnitude of the cochlear reflectance varies relatively slowly with frequency, showing none of the spectral periodicity seen in measurements of the ear canal pressure. Thus to a first approximation these experimental results agree with those from the idealised models presented earlier.

The ways in which measured SFOAEs depart from the idealisation are also worth noting. Three main ways can be identified (e.g., Shera and Zweig, 1993a). Firstly, they are nonlinear, except perhaps at very low stimulus levels (section 3.5). Secondly, the cochlear reflectance is not perfectly constant, but instead shows a slow, random fluctuation with frequency. Thirdly, the middle ear is far from reflectionless, leading to a more complicated series of multiple reflections than that seen in fig. 2.5. This is discussed further in part II.

2.8 Alternative Definitions of SFOAEs

Confusingly, in the literature, the term 'SFOAE spectrum' does not have a consistent meaning. In some papers (e.g., Talmadge *et al.*, 1998), it means the spectrum of the total ear canal pressure response (i.e., stimulus signal plus the echo-like component), which then shows a strong pattern of ripples in its magnitude. This corresponds to panels (i) and (j) in figs. 2.4 and 2.5. Elsewhere, however, the term 'SFOAE' is used to mean the spectrum of the echo-like component alone, which does not necessarily show the same regular pattern of ripples in its magnitude (e.g., Kemp and Chum, 1980a). This corresponds to panels (f)

and (g) in figs. 2.4 and 2.5. In this thesis, the term 'SFOAE' refers to the echo-like component only, whilst the term 'ear canal pressure spectrum' is used to refer to the total spectrum recorded by the microphone. Thus the ear canal pressure spectrum is the (complex) sum of the stimulus component and the SFOAE component.

2.9 Measurement Methods for SFOAEs

In the idealised spectra shown in figs. 2.4(i) and 2.5(i), the ear canal pressure signal can be easily split into the stimulus and SFOAE components, because the stimulus component is a known signal (in this case entirely uniform). The situation is less simple in the case of real OAEs. Two issues are addressed in this section. Firstly how to define the stimulus component conceptually and secondly how to measure it.

Recall that in fig. 2.5(a) the stimulus component was defined as the acoustic pressure due to the initial forward going wave, and the OAE component was then defined simply the remainder of the ear canal signal. Because the idealised middle ear was non-reflecting, this was a convenient definition, since the resulting OAE component originated entirely in the cochlea. However, in reality, the middle ear is reflecting. This means that multiple reflections will occur in the ear canal, whether or not the cochlea has a non-zero reflectance. Therefore, it is useful to modify the definition of the stimulus component from that used previously, such that it excludes signals originating from backward TW in the cochlea, but includes reflections from the eardrum[7]. The stimulus component is therefore redefined as the pressure signal that would have arisen had the cochlea been entirely non-reflecting[8].

Using this new definition, the typical characteristics of the stimulus component will now be considered. Assuming the eardrum is approximately 10 mm from the loudspeaker, the forward going wave from the loudspeaker will be partially reflected after about 0.03 ms. Thus, on striking the eardrum, the initially forward going wave splits into two components. The first is a transmitted wave which goes on to enter the cochlea[9] where it

[7] These reflections are dependent on the properties of the passive (or reflectionless) cochlea, as these affect the eardrum impedance.

[8] The question of what is meant by a 'non-reflecting' or 'reflectionless' cochlea will be addressed more fully in part II of this thesis. Here it is simply noted reflectionless means that a forward cochlear TW propagating along the BM does not give rise to a backward TW.

[9] For simplicity in this explanation, the ossicular chain has been treated as rigid linkages. In reality there will be reflections not only from the eardrum, but also from other points in the ossicular chain. However, since these are separated by extremely short delays, they may be treated as a single reflection.

generates a forward TW on the BM which in turn elicits an OAE. The second is a reflected wave which then reverberates in the ear canal, until it decays away typically over about 3 ms. Each time a forward waves strikes the eardrum it further splits into two components. The stimulus component is equal to the initial forward wave, plus the first reflection from the eardrum, plus subsequent reflections between eardrum and ear canal termination. In fact, because of the short delays between these reflections, all these components add up constructively for the frequencies that are of interest to us here (i.e., less than 4 kHz), such that the stimulus wave is larger than the initial forward going wave component. In addition to the effect of reflections within the ear canal, the transient response of the loudspeaker also influences the stimulus component for any given voltage input signal. Thus overall, the stimulus component depends on the dynamic properties of the loudspeaker, the ear canal termination, the ear canal itself, the middle ear and the passive (or reflectionless) cochlea. Since many of these properties are unknown, it is impossible to calculate the stimulus component just from knowledge of the voltage input to the loudspeaker. Practical methods of estimating the stimulus component are now discussed.

First consider the measurement of TEOAEs. For an ideal voltage click, the stimulus component rings on for around 3 ms, obscuring any early TEOAE components. However, because the OAE delays are typically much longer than the decay time for the stimulus (it is common for TEOAEs to remain measurable up to and beyond 25 ms (Probst *et al.*, 1991), the majority of the TEOAE remains unobscured. A common method of separating TEOAEs from the stimulus is simply to discard the first few milliseconds of the ear canal pressure signal.

This same situation can also be considered in frequency domain (cf., fig. 2.5(e)-(g)). For a pure tone input, the effect of the reverberation is to cause an amplitude and phase shift of the stimulus pressure component relative to the voltage input. However, this shift varies only very slowly with frequency, because of the very short time delays. For example, the ripple pattern associated with the ear canal reverberation discussed above would have a peak-to-peak frequency spacing of $1 / 0.06$ ms $\approx$ 17 kHz. This compares with a ripple spacing seen for SFOAEs of around 0.06 kHz at 1 kHz. Thus, in the frequency domain, the ear canal ripple is not superimposed onto a perfectly uniform background line, as it was in the idealised case in fig. 2.5(i). Instead the background slowly fluctuates with frequency, due to the frequency response of the passive system. However, because this fluctuation in the frequency domain is relatively slow, the phasor diagram in fig. 2.2 showing a rapidly rotating OAE component plus a slowly varying stimulus component is still a useful one. One method of estimating the SFOAE component uses the fact that the

spectral ripples due to the OAE component have a much shorter frequency interval than the slow variation due to the stimulus. A spectral smoothing technique which removes the ripples, leaving the underlying stimulus component is reported by Shera and Zweig (1993a). Having obtained the stimulus component, it is then subtracted from the ear canal pressure signal to give the OAE component. (Note that all three are complex.) In fact, smoothing in the frequency domain is equivalent to windowing (or truncation) in the time domain. Thus this technique is equivalent to inverse Fourier transforming the ear canal pressure, rejecting the early part (as for TEOAEs above), and then Fourier transforming back again. It should be borne in mind that the SFOAE is not obtained from a linear system and therefore the inverse Fourier transform of the SFOAE is not the same as the TEOAE (it may not even be perfectly causal). The optimum truncation may therefore be different for the two classes of OAE.

Two other practical methods of obtaining the stimulus component are the self-suppression and two-tone suppression techniques. These make use of nonlinear acoustic suppression, where it is assumed that measured suppressed cochlear responses are approximately equal to passive cochlear responses. The self-suppression method makes use of the compressive nonlinearity. Here an estimate of the slowly varying stimulus spectrum is obtained by measuring the (complex) ear canal pressure at a high stimulus level (say 70 dB SPL), where the ripple pattern is relatively weak (Kemp and Chum, 1980a). The two-tone suppression method estimates the slowly varying stimulus spectrum by introducing a high level suppressor tone to suppress the ripple (Kemp and Brown, 1983; Kemp *et al.*, 1990). Having obtained an estimate of the (complex) stimulus spectrum alone, it is then subtracted from the (complex) unsuppressed ear canal spectrum, yielding the SFOAE component alone. Note that analogous methods relying on compressive nonlinearity are also used for the measurement of TEOAEs to remove so called click artefacts (Probst *et al.*, 1991; Molenaar *et al.*, 2000).

In general, these three methods appear to yield very similar results, indicating that the spectrally smoothed ear canal pressure is approximately linear[10] and therefore unaffected by level changes or suppressor tones (Shera and Zweig, 1993a). The correspondence of the spectral smoothing technique and the acoustic suppression technique is consistent with theoretical predictions, discussed in part II, that the passive component is essentially both linear and of short latency, while the active component is both nonlinear and predominantly of long latency. In this project the two-tone suppression method has

[10] A theoretical analysis of the validity of this approximation and of the differences between these three methods is given by Talmadge *et al.* (2000).

been adopted.

Note that both the SFOAE component and the stimulus component are affected by the acoustic impedance of the OAE probe that that is sealed into the ear canal. This is because the reflections (whether passive or active) reverberate in the ear canal, which is closed at its outer end by the OAE probe. It is possible to measure the impedance of the probe using hard walled cavities, and therefore to obtain a (level dependent) measure of the impedance looking into the ear canal. This is then independent of the probe characteristics (Jurzita and Hemmert, 1992). An alternative approach to eliminating the effect of the probe characteristics in which the probe reflectance is measured is reported by Keefe (1997). However, provided that the source impedance of the loudspeaker is high, the effect of the probe is only secondary and there is little advantage in using these methods in this thesis.

2.10 Experimental Data on the Group Delay of SFOAEs

In the idealised ear in figs 2.4 and 2.5 it was shown that the logarithmic variation of phase gave rise to a $1/f$ variation of group delay. This is illustrated in figs. 2.6 a and b (ignoring the ripples seen in fig 2.6b). An approximate $1/f$ variation has been found in experimental measurements of group delay. Averaging across seven ears, Kemp and Chum (1980a) found the group delay varied with frequency from about 18 ms at 0.6 kHz to about 10 ms at 1.4 kHz. Similar results were reported by Wilson (1980a) who used tone burst stimuli of 5 or 8 cycles, and estimated the latency of the envelope in the time domain.

The ripples in the group delay of the idealised data fig. 2.6(b) illustrate one of the properties of the group delay of a signal. Recall that the group delay can be interpreted as the delay of the envelope of a tone burst response to tone burst excitation. This interpretation must be applied carefully in cases where multiple reflections are present, such as those illustrated in figs 2.3 and 2.5. In these cases, a tone burst input would give rise to multiple tone bursts in the response, each with its own delay. However, the calculation of group delay yields only a single value of delay for any given frequency. To interpret the group delay correctly in these cases, we must consider the excitation to be a very long tone burst, such that its envelope is much longer than the decay time for the train of multiple reflection components. Then in the response, these multiple components interfere with each other, and form one single tone burst whose envelope latency can be determined. The result of this is that the interference between components causes ripples in the group delay as seen in fig. 2.6. Note, however, that when averaged over frequency such that the ripples disappear, the group delay gives the delay of the first (and strongest)

component, rather than a weighted average of the delays of all the components. As is discussed later, this does not always give a measure of delay which is physically useful. The group delays reported in the literature are commonly estimated by fitting a straight line to the phase curve over a significant frequency interval, thus effectively performing a form of frequency averaging which will remove any ripples. This also avoids the problem that any point estimate of the slope would be highly sensitive to the signal to noise ratio. Kemp and Chum (1980a) estimated the group delay by averaging over a frequency region that gave a monotonic drop in phase of about half a cycle. Kemp and Brown (1983) used a frequency interval of about half an octave.

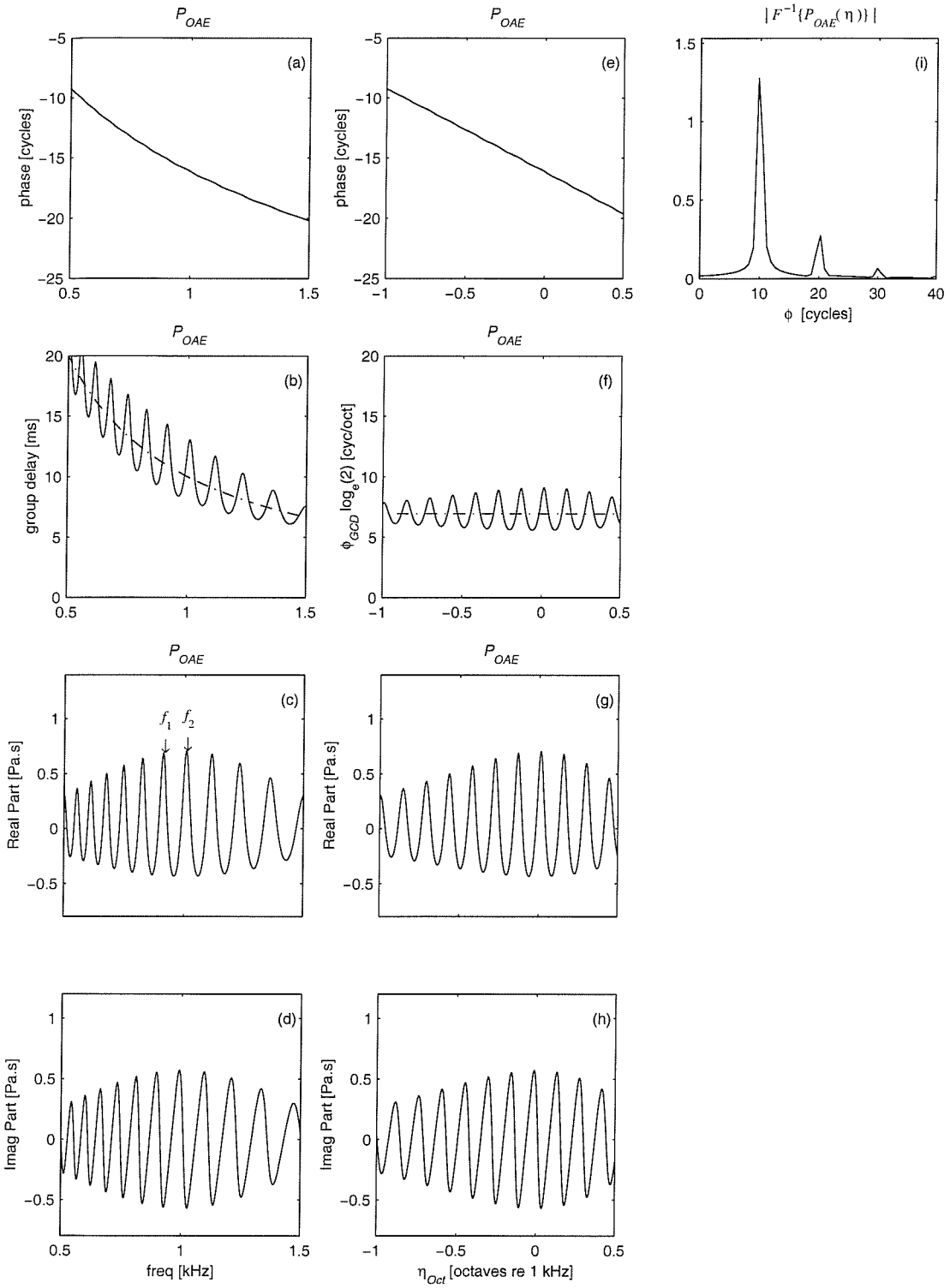


Fig 2.6. Signal processing of the idealised OAE pressure component shown in fig 1.4. Panel (a) shows the phase of the frequency spectrum of the OAE component, given by $\theta_1(f) \approx -A \log_e(f/f_{ref})$. Panel (b) shows the group delay, defined by $\tau_{GD}(f) \equiv -(1/2\pi)\, d\theta_1(f)/df$. This approximates to $A/(2\pi f)$, shown by the dashed line. Panel (c) shows the real part of the frequency spectrum, with two adjacent spectral peaks marked by f_1 and f_2. Panel (d) shows the imaginary part of the frequency spectrum. Panels (e), (g) and (h) show the functions in (a), (c) and (d) evaluated against the logarithmic frequency variable, $\eta_{Oct} \equiv \log_2(f/1\text{ kHz})$. Panel (f) shows $-(1/2\pi)\, d\theta_1(\eta_{Oct})/d\eta_{Oct}$ which equals the group cycle delay, multiplied by $\log_e(2)$. This evaluates to: $\log_e(2) \times \varphi_{GCD}(\eta_{Oct}) \approx A \log_e(2)/(2\pi) \approx 6.9$ cycles/octave (shown by the dashed line). Panel (i) shows the magnitude of the φ-spectrum, defined as the inverse Fourier transform of $P_{OAE}(\eta)$.

2.11 Spectral Periodicity of SFOAEs

Rather than measure the group delay, many researchers have characterised the ripple pattern due to the SFOAE in terms of the ripple spacing in the frequency domain. The "ripple frequency interval", Δf, is defined simply as:

$$\Delta f \equiv f_2 - f_1 \qquad [2.8]$$

where f_1 and f_2 are the frequencies of adjacent ripple peaks, as shown in fig. 2.6(c) for the idealised ear[11].

As discussed in section 2.5, the ripple spacing is, to a first approximation, proportional to the frequency (Kemp and Chum, 1980a; Wilson, 1980a; Zwicker and Schloth, 1984; Dallmayr, 1987; Zweig and Shera, 1995). Consequently, it is common to quantify the ripple spacing for a pair of ripple peaks using a quantity called here the "single-ripple spectral periodicity", Ψ , defined as the ripple frequency interval divided by the geometric mean frequency, f_{GMF}, of the two peaks:

$$\Psi \equiv \frac{\Delta f}{f_{GMF}} \equiv \frac{f_2 - f_1}{\sqrt{f_2 f_1}} \qquad [2.9]$$

(e.g., Zweig and Shera, 1995). SFOAE measurements show that the spectral periodicity, Ψ, is roughly constant over a frequency range from 0.5 to 4 kHz, and is equal to about 1/15, or 7% (Zwicker and Schloth, 1984; Dallmayr, 1987; Zweig and Shera, 1995). Note that the invariance of Ψ is only a first approximation – several authors report that Ψ reduces somewhat with increasing frequency (e.g., Zweig and Shera, 1995; p. 2036).

It is shown in the following analysis that a constant single-ripple spectral periodicity corresponds to equal ripple spacing on a logarithmic frequency scale. Two logarithmic scales are used in this thesis. The first is the octave scale, denoted η_{Oct}, which is convenient due its familiarity, and the second is the natural logarithm of frequency, denoted η, which has theoretical advantages that will become apparent later. These are defined by:

$$\eta_{Oct}(f) \equiv \log_2(f / f_{Ref}) \qquad f_{Ref} \equiv 1\,\text{kHz} \qquad [2.10]$$

and

$$\eta\ (f) \equiv -\log_e(f / f_0) \qquad [2.11]$$

[11] The ripple spacing in the real part of the OAE spectrum is the same as that in the ear canal pressure magnitude.

where (following Zweig and Shera, 1995) f_0 is defined as the characteristic frequency at the very base of the cochlea. The minus sign in [2.11] is included for compatability with Zweig and Shera (1995), and leads to simplifications later on.

Given the two frequencies, f_1 and f_2, used to define the periodicity, we can define the "ripple natural logarithmic frequency interval", $\Delta\eta$, as $\eta(f_1) - \eta(f_2)$. Then manipulating equations [2.9] and [2.11], we obtain the following relationship between $\Delta\eta$ and the periodicity, Ψ:

$$\Delta\eta = 2\log_e\left(\Psi/2 + \sqrt{1+\Psi^2/4}\right)$$

or [2.12]

$$\Delta\eta \approx \Psi \quad \text{for } \Psi << 1$$

The approximation in [2.12] is a very close one for typical Ψ values, which are around 1/15. The corresponding relationship for the "ripple octave interval", $\Delta\eta_{Oct}$, defined as $\eta_{Oct}(f_2) - \eta_{Oct}(f_1)$ is:

$$\Delta\eta_{Oct} = \frac{\Delta\eta}{\log_e(2)} \approx 1.443\Psi \quad \text{for } \Psi << 1 \qquad [2.13]$$

Thus, [2.12] and [2.13] show that a constant periodicity implies a constant ripple spacing on a logarithmic frequency scale. For the idealised OAE, fig 2.6 shows how transforming to an octave scale, η_{Oct}, leads to a linear phase curve in panel (e), and equal ripple spacing in the real and imaginary parts of the Fourier transform of the OAE in panels (g) and (h).

The linear phase curve, in fig. 2.6(e), suggests a second way of characterising the ripple periodicity. Consider the quantity which we will call the "group cycle delay", denoted $\phi_{GCD}(\eta)$, and which is defined in a similar way to the group delay in equation [2.6], but now using the logarithmic frequency. Thus we define:

$$\phi_{GCD}(\eta) \equiv -\frac{1}{2\pi}\frac{d\theta}{d\eta} = -\frac{1}{2\pi}\frac{d\theta}{df}\frac{df}{d\eta} = -f\frac{1}{2\pi}\frac{d\theta}{df} = f\tau_{GD}(f) \qquad [2.14]$$

$$\theta \equiv \arg(\boldsymbol{P}_{OAE})$$

There is a useful physical interpretation of ϕ_{GCD} which follows from [2.14], and which has been reported by several authors (Wilson, 1980a; Kemp, 1986; Zweig and Shera, 1995). Recall that the ordinary group delay indicates the delay in units of time of the envelope of a tone burst signal. Multiplying the ordinary group delay of the tone burst by its carrier frequency, as in [2.14], converts the delay in units of time to a delay measured as a number of cycles of the carrier signal. Note that this interpretation holds for any arbitrary frequency response, not just for the OAE signals analysed here. However, for OAE signals, ϕ_{GCD} is roughly constant over frequency. This follows from the definition of $\Delta\eta$ in [2.12] , which implies that:

$$\phi_{GCD} = 1/\Delta\eta \approx 1/\Psi \qquad [2.15]$$

for an idealised cochlear reflectance. For measured OAE signals, the fact that Ψ, and therefore ϕ_{GCD}, are found to be roughly independent of frequency means that the delay between the peaks in the envelopes of the input and output tone bursts is roughly a constant number of carrier wave cycles. The number of cycles equals the reciprocal of the periodicity, and is thus typically between 12 and 15 cycles (Wilson, 1980a; Kemp, 1986; Zweig and Shera, 1995).

Fig 2.6(f) shows the gradient of the phase curve on the η_{Oct} scale for the idealised ear, given in units of cycles per octave. This is proportional to $\phi_{GCD}(\eta)$. The ripples seen in this curve arise from the presence of multiple reflections. A single reflection would give the constant value shown by the dashed line. Note that in this idealisation, we have defined $\theta = \theta_1 \approx -A \log_e (f/f_{Ref})$, which gives $\phi_{GCD} \approx A/2\pi$. The actual value in the fig. 2.6 is 10 cycles, corresponding to a periodicity of 0.1, or about 6.93 cycles per octave. (This periodicity is somewhat higher than the value of 0.07 commonly found in the literature.)

A third method that has been used to characterise periodicity is to apply the Fourier transform to the spectrum of the SFOAE, after it has been transformed onto the logarithmic frequency axis as in figs. 2.6 (e)-(h) (e.g., Zweig and Shera, 1995; Lutman and Deeks; 1999). A simple justification for this is that this transformed SFOAE spectrum has a roughly periodic ripple pattern, which will show up as lines in the Fourier transform. A better physical interpretation of this can be gained by noting that the spectrum of the idealised OAE in the logarithmic frequency domain shown in figs. 2.6 (g) and (h) resembles the spectrum in the ordinary frequency domain of the echo in the multiple reflection case shown in figs 2.3 (f) and (g). Now taking the inverse Fourier transform of this latter spectrum yields the time domain signal in fig. 2.3 (e). Similarly, taking the inverse Fourier transform of figs 2.5 (g) and (h) would yield a series of lines in the "$1/\eta_{Oct}$-domain". A closely related function is illustrated in fig. 2.5 (i), which represents the signal in the "$1/\eta$-domain" rather than the "$1/\eta_{Oct}$-domain". This figure was obtained by first replotting the spectrum in figs 2.5 (g) and (h) against the η variable, and then taking the forward Fourier transform[12]. Following Zweig and Shera (1995), the

[12] It is a general result that, for any arbitrary signal, $x(t)$, the forward Fourier transform of $x(t)$ yields the same result as the inverse Fourier transform of $x(-t)$ (where both transforms are taken w.r.t. the t variable). It follows from the definitions of η_{Oct} and η, that the inverse transform from the η_{Oct} domain is simply a scaled version of the forward transform from the η domain.

$1/\eta$-domain will henceforth be referred to as the ϕ-domain. By analogy with fig. 2.3 (e), the value of ϕ for a given line in the ϕ-domain indicates the delay of that OAE component, but measured in units of cycles rather than time. The dominant line in the ϕ-domain lies at the ϕ_{GCD} (which is $1/\Psi$), which is 10 cycles in fig. 2.5. The ϕ-domain then represents the OAE response to a click in which the $1/f$ – dispersion has been "undone" by an appropriate transformation, such that the frequency components all coincide to form a single impulse (or a series of impulses in the case of multiple reflections).

To summarise, the SFOAE periodicity arises from the time delay between the stimulus signal and the OAE signal. For the idealised SFOAE considered here, three ways of characterising the SFOAE periodicity have been introduced: the single-ripple periodicity, Ψ; the group cycle delay, ϕ_{GCD}; and the peak in the ϕ-domain. The approximate independence of frequency of the periodicity corresponds to a dispersive group delay, τ_{GD}, which shows a $1/f$ frequency dependence. This implies a constant group cycle delay, ϕ_{GCD}, equal to $1/\Psi$. Transforming the SFOAE spectrum onto a logarithmic frequency scale, η, linearises the phase relationship, thereby "undoing" the effect of the dispersion. The group cycle delay can also be estimated from the inverse Fourier transform of the SFOAE spectrum from the η-domain to the ϕ-domain. This also reveals the presence of multiple reflection components in the SFOAE. Note that these three characterisations of periodicity are only exactly equivalent for the idealised SFOAEs generated in figs. 2.4 and 2.5. In real SFOAEs they may still be closely related to each other, but cannot be used interchangeably.

A final point to note is that instead of characterising the SFOAE ripple pattern using a logarithmic frequency transformation, a measure based on critical bandwidth is sometimes reported (e.g., Zwicker and Schloth, 1987; Dallmayr, 1987; Zwicker, 1988). Such a measure typically gives the interpeak frequency spacing as 0.4 Bark. This measure is based on Zwicker and Schloth's observation (1987), that the shape of the graph of ripple frequency interval, Δf, against centre frequency, f_{GMF}, more closely follows a curve of critical band width against frequency than a straight line. This measure has not been adopted for this project, since the stimulus frequencies of interest are all above 1 kHz where the Bark scale differs little from a purely logarithmic frequency scale.

2.12 Practical Methods of Quantifying the SFOAE Periodicity

In this project, it is necessary to obtain a measure of the average periodicity seen in the SFOAE signal over a certain stimulus frequency range. Measured SFOAE signals

differ substantially from the idealised OAE signals shown in fig. 2.6, making estimates of the SFOAE more problematic. Two differences are important here. Firstly, the periodicity varies systematically with frequency (Zweig and Shera, 1995). Secondly, the ripple pattern is much less regular than the virtually pure sinusoid seen in the idealisation. This irregularity often leads to frequency regions where the ripple pattern disappears below the measurement noise floor. Thus rather than the SFOAE showing one clear periodicity, it often appears that a distribution of periodicities is present in the signal. To address this problem, more realistic phenomenological models of the SFOAE signal will be discussed more fully in parts II and III. In this section, some of the different methods for estimating the periodicity that have been reported in the literature are reviewed.

In the literature, several different methods have been used to quantify the periodicity of a measured SFOAE frequency spectrum. A common method is to directly measure the frequency intervals between adjacent peaks in the ripples pattern seen in the magnitude of the ear canal pressure frequency spectrum. If the peak-to-peak interval, Δf, is plotted against the peak-to-peak centre frequency then the points lie roughly on a straight line whose gradient gives a direct measure of the average single-ripple periodicity over the measured frequency range (Dallmayr, 1987; Zwicker and Schloth, 1987; Zwicker, 1990; Zwicker and Peisl, 1990; Lonsbury-Martin *et al.*, 1990). A problem with this method is that the peaks in the spectrum are not always clear. Therefore, some form of peak identification algorithm must be used which ignores very small ripples which may be due to noise.

A second method of quantifying the periodicity is to measure the slope of the phase of the SFOAE frequency spectrum. This gives a measure of the group delay (in units of time) of the SFOAE (Kemp and Chum, 1980a; Wilson, 1980a; Kemp and Brown, 1983). For the idealised system, the group delay, $\tau_{GD}(f)$, and the single-ripple periodicity, Ψ, are related via equations [2.14] and [2.15]. Thus, fitting a straight line to a plot of τ_{GD} against $1/f$ will yield a measure of periodicity (cf., fig. 2.6b). A related method of obtaining a periodicity value is to plot the phase of the SFOAE frequency spectrum against the logarithm of stimulus frequency, η, in order to straighten out the phase curve (cf., fig. 2.6f). The gradient of the best-fit straight line then directly yields a measure of the group cycle delay, $\phi_{GCD}(\eta)$, and therefore, from [2.15], a measure of periodicity. Wilson (1980a) reported a group cycle delay equal to about 15 cycles, over a frequency range of 0.5 to 5 kHz. Kemp and Brown (1983) demonstrated a group delay in humans varying with stimulus level from about 18 cycles down to about 8 cycles for a stimulus range from 10 to 70 dB SPL.

One problem with the group delay (or group cycle delay) method (at least for the purposes of this project) is that the phase of the SFOAE becomes hard to measure when the magnitude is small. A second problem is that any averaging over frequency takes no account of the SFOAE magnitude. Thus contributions from regions where the SFOAE is weak (and therefore where the group delay is unreliable) receive the same weight as contributions from regions where the SFOAE is strong. A third and more serious problem is that, unlike the idealised SFOAEs in fig 2.6g, real SFOAE signals do not resemble a sine wave, but instead appear as a "bandpass" signal[13]. This means that a number of components of different periodicity are present in a single SFOAE spectrum. In the time domain, this can be interpreted as multiple bursts of OAEs with the same carrier frequency, but with different envelope delays[14]. Since the group delay can only yield one overall envelope delay per carrier frequency, some sort of averaging must occur. As a consequence, it turns out that the way in which the group delay quantifies the periodicity does not correspond very well to the physical quantities in Shera and Zweig's theory. This will be discussed more fully later

A third method of characterising the periodicity is reported by Zweig and Shera (1995). Here, the complex SFOAE pressure is first plotted as a function of the logarithm of frequency, η, such that the ripple pattern appears roughly periodic (for constant periodicity). The resulting waveform is then (either forward or inverse) Fourier transformed and the location of the peak of the magnitude of this transform is used as a measure of the average periodicity (cf., fig. 2.6(i)). A similar method was used by Lutman and Deeks (1999) who plotted the amplitude of the ear canal pressure (rather than the complex SFOAE) on a log frequency scale before Fourier transforming. (A related method of processing TEOAE spectra is reported by Wit *et al.* (1994), but they omit the logarithmic transformation thus leaving the effect of dispersion unchanged from that seen in the time domain.) These methods have the practical advantage that areas where the SFOAE is weak (and less reliable) contribute less strongly to the average periodicity. More importantly, Zweig and Shera's method yields a periodicity measure which can be related quantitatively to other parameters of interest via their own theory of SFOAE generation. For this reason, a periodicity measure based on this approach has been developed for this project. These methods are discussed further in part II.

---

[13] Bandpass here means that the SFOAE plotted against the η variable resembles the more familiar bandpass time series, obtained by passing white noise through a bandpass filter.

[14] Consider, for example, the result of adding two idealised OAE signals, each resembling that in fig 2.4, but each with a slightly different value of A. See also the discussion in section 2.10.

2.13 Linearity and Causality of SFOAE Signals

The idealised OAEs in figs. 2.3, 2.4 and 2.6 were generated from a linear, causal system. The question of whether real measured SFOAEs show either of these properties was addressed by Shera and Zweig (1993a). They found that the amplitude of SFOAEs grew linearly with stimulus levels for levels between 0 dB SL and 5 dB SL. This suggests that a low level linear regime exists for SFOAEs which is important since it allows results from a linear theory to be applied. As already discussed, it is known that at higher levels SFOAEs are far from linear. Instead they show a compressive nonlinearity which greatly complicates both the theory and any signal processing.

Shera and Zweig (1993a) also found that the SFOAEs in the linear regime were consistent with the response of a causal system, indicating that the OAE response always occurs after the evoking stimulus. This result contradicted earlier results published by Zwicker and Schloth (1984) which had shown SFOAE signal characteristics which where inconsistent with those of a causal system. This had prompted speculation that, for a pure tone input, the brain might anticipate the input signal, and modify the cochlea[15] (Shera and Zweig, 1993a) such that the SFOAE response appeared acausal.

For any given frequency response function, a simple way to test for causality (for a linear system) is to inverse Fourier transform the signal to give the time domain impulse response, and then to check that this response is approximately zero for all negative times. An equivalent way of doing this entirely within the frequency domain is to use the Hilbert transform. It is useful to consider this method here as it reveals an important property of SFOAEs that will be used later on in parts II and III. This method is best understood by first noting that any time series $x(t)$, can be split into the sum of even and odd components, $x_E(t)$ and $x_O(t)$:

$$\begin{aligned} x(t) &= x_E(t) + x_O(t) \\ x_E(t) &\equiv \tfrac{1}{2}x(t) + \tfrac{1}{2}x(-t) \\ x_O(t) &\equiv \tfrac{1}{2}x(t) - \tfrac{1}{2}x(-t) \end{aligned} \qquad [2.16]$$

This is useful since the Fourier transform of an even signal is purely real whilst that of an odd function is purely imaginary. Thus the real and imaginary parts of the Fourier transform of $x(t)$ are given by the Fourier transform of the even and odd components in

[15] Such a system would not be truly acausal in that the evoking stimulus must still preceed any response. However, the system might quickly adapt nonlinearly to a prolonged stimulus in such a way that, after having adapted, the system appeared to behave linearly but acausally. In Shera and Zweig's data (1993a), no acausality was found.

[2.16]. Now for a causal system it can easily be shown that the odd and even part are not independent, but are related by:

$$x_O(t) \equiv x_E(t)\,\mathrm{sgn}(t)$$

where

$$\mathrm{sgn}(t) = -1 \quad \text{for } t < 0 \qquad +1 \quad \text{for } t \geq 0 \qquad [2.17]$$

$$x_E(t) = \tfrac{1}{2}x(t) \quad \text{for } t \geq 0$$

This means that the real and imaginary parts of the Fourier transform are also not independent, but are related via a convolution with the Fourier transform of signum function, $\mathrm{sgn}(t)$. This latter operation can be shown to be identical to the Hilbert transform (Randall, 1987). Thus:

$$\begin{aligned} X_{\mathrm{Re}}(f) &\equiv \mathrm{Re}\{F\{x(t)\}\} = F\{x_E(t)\} \\ X_{\mathrm{Im}}(f) &\equiv \mathrm{Im}\{F\{x(t)\}\} = F\{x_O(t)\}/i \\ &= X_{\mathrm{Re}}(f) \otimes F\{\mathrm{sgn}(t)\}/i \\ &= -\mathrm{Hilb}\{X_{\mathrm{Re}}(f)\} \end{aligned} \qquad [2.18]$$

where X_{Re} and X_{Im} denote the real and imaginary parts of the Fourier transform of $x(t)$; $F\{\ \}$ denotes the Fourier transform; $\otimes$ denotes convolution and Hilb $\{\ \}$ denotes the Hilbert transform. It can be shown that the Hilbert transform operation acts as an all-pass filter which delays the output signal by 90° relative to the input. For example, the Hilbert transform of a cosine wave is a sine wave of the same frequency. This behaviour can be seen in fig. 2.6(c) and (d) where the imaginary part lags the real part by 90°. Shera and Zweig (1993a) showed that, for their measurements, the real and imaginary parts did indeed form a Hilbert transform pair, implying that the system relating the SFOAE to the evoking stimulus is causal. In fact they showed that the system was not only causal, but also minimum phase, which is characteristic that is commonly found in physical systems. For a causal minimum phase system, not only the real and imaginary parts, but also the magnitude and phase of the transfer function are related via the Hilbert transform (Randall, 1987).

2.14 Relationships between SFOAEs and other Phenomena in Hearing

In the previous sections, some of the characteristics of measured SFOAE signals were discussed. Some of these characteristics are also seen in other phenomena in hearing. The spectral periodicity seen in SFOAEs is also seen in the spectra of click- and tone burst evoked TEOAEs, in the fine-structure of DPOAEs, in the distribution of the frequency

spacing between SOAEs, and in the audiometric microstructure (Elliot, 1958; Kemp, 1979; Probst *et al.*, 1986; Schloth, 1983; Zwicker and Schloth, 1984; Dallmayr, 1987; He and Schmiedt, 1996, 1997; Kapadia and Lutman, 1999; Lutman and Deeks; 1999). Since all these phenomena appear to be related, the development of a theory of SFOAE generation must form part of a more general theory of cochlear mechanics which accounts for all these phenomena. Furthermore, effects that show up in SFOAE measurements may also appear in these related phenomena. Thus in the following sections, the discussion will not be limited to SFOAEs.

SFOAEs also share several general characteristics with the other classes of OAE, with certain psychoacoustic phenomena, and with cochlear mechanical responses measured in animals. These characteristics include saturation at high levels, two-tone suppression effects, and vulnerability to ototoxic drugs such as aspirin (e.g., Long and Tubis, 1988; Karlsson *et al.*, 1991; and reviewed in Probst *et al.*, 1991). As with other OAEs, SFOAEs are strongly related to hearing threshold levels, and are not generally measurable in subjects with a hearing loss exceeding about 30 dB. These phenomena are relevant since they involve changes in TW shape.

2.15 Effect of Stimulus Level on OAEs (Self-suppression)

Several researchers have investigated the effect on SFOAEs and TEOAEs of increasing stimulus level. For pure tones below about 5 dB above the threshold of hearing, SFOAEs grow approximately linearly with stimulus level (Shera and Zweig, 1993a). However, as the stimulus level is increased further, they show compressive nonlinearity (Kemp and Chum, 1980a; Zwicker and Schloth, 1984; Zwicker, 1990; Dallmayr; 1987). Compressive nonlinearity is of particular importance to this thesis, since both cochlear models (e.g., Kanis and de Boer, 1993b) and direct measurements (e.g., Rhode, 1971) indicate that, as the stimulus level increases, the peak of the TW envelope becomes broader, and the TW wavelength[16] in the peak region increases. According to Shera and Zweig's theory, these changes should be accompanied by an increase in spectral

[16] From theoretical considerations of their cochlear model, which exhibits the minimum phase property, Zweig and Shera (1995) demonstrate that the TW wavelength and the width of the envelope are not independent.

periodicity of SFOAEs (or equivalently, a reduction in the latency of TEOAEs)[17]. In addition, the 'bandwidth' of the periodicity should reduce with stimulus level (see part II).

Several authors have measured SFOAEs at different stimulus levels, without reporting any effect of level on spectral periodicity (e.g., Kemp and Chum, 1980a; Zwicker and Schloth, 1984; Zwicker, 1990). Dallmayr (1987, fig. 8b) explicitly stated that stimulus level did not affect the SFOAE phase spectrum (related to the spectral periodicity). These results suggest that the effects predicted by Shera and Zweig's theory are either absent or difficult to detect. In contrast, Zweig and Shera (1995, fig. 13) remark that they have recorded changes in the spectral periodicity of SFOAEs with level which match the predictions of their cochlear models. However, no measured data were presented and no quantitative analysis was attempted.

Prior to the development of Shera and Zweig's theory (1993b), Kemp and Brown (1983) measured SFOAEs in human and gerbil ears at various stimulus levels. They measured the group delay (i.e., the slope of the phase spectrum) and reported a reduction in latency with increasing level, as is now predicted by Shera and Zweig's theory. However, no detailed investigation of the effect was performed and nor was any attempt made to compare the results with theoretical predictions from cochlear models. A further point to note is that the group delay method of detecting level effects is not well suited to testing Shera and Zweig's theory (see part II).

Wilson (1980a) measured group delays at a few frequencies using tone burst stimuli, rather than pure tones. He also reported a reduction in group delay with increasing stimulus level. However, quantitative results were not reported. A comprehensive investigation of tone-burst evoked OAEs, in which latency was measured using a correlation technique, was reported by Norton and Neely (1987). They too reported a reduction in latency with increasing stimulus level. They interpreted their results in terms of the speed of the TW (Neely *et al.*, 1988), and suggested that the latency was consistent with measures of ABR latency made at comparable stimulus levels. This interpretation is based on the assumption that the OAE latency depends on the TW propagation time to and from a single place-fixed reflection site. According to Shera and Zweig's theory, this is an oversimplification which nevertheless leads to reasonably accurate predictions. This will be discussed further in section 3.3.3 and in part II.

---

[17] Recently, Talmadge *et al.* (2000) have presented an analysis of Shera and Zweig's theory based on nonlinear cochlear models which suggests that this simple qualitative prediction cannot always be made. This is because, in some cases, an additional effect of nonlinearity is to disrupt the simple relationship between TW shape and SFOAE periodicity (see section 3.5 for a fuller discussion).

Some researchers have looked for a direct relationship between the amplitude of OAEs and the sharpness of the auditory filter (measured psychophysically), both of which are thought to be related to the shape of the TW. Leeuw and Dreschler (1998) found significant correlation between these two quantities within subjects on varying the stimulus level. Micheyl and Collet (1994) found a correlation looking across subjects at a constant stimulus level. However these experiments examined the relationship of TW shape to SFOAE amplitude, rather than to SFOAE periodicity, and therefore are not directly related to the aims of this project. Avan *et al.* (2000) looked at differences in the periodicity of TEOAE spectra between normally hearing adults and adults with mild noise-induced hearing loss. It would be expected from Shera and Zweig's theory that the latter group would show increased periodicity due to the reduction in sharpness of the TW envelope that accompanies cochlear hearing loss. However, in practice, the periodicity proved difficult to measure in this group.

In summary it appears that changes in SFOAE periodicity with stimulus level predicted by Shera and Zweig's theory are either absent or difficult to detect with the signal processing methods previously used. Though some experimental evidence that supports this theoretical prediction has been reported (Wilson, 1980a; Kemp and Brown; 1983; Zweig and Shera; 1995), no systematic attempt to relate measurements and theory has been published. Also, no attempt to measure the effect of stimulus level on the bandwidth of the spectral periodicity, rather than the centre value of the periodicity, has previously been reported. These are two of the effects that are investigated in this thesis.

2.16 Effect of Suppressor Tones on OAEs

As with stimulus level, the effect of introducing a suppressor tone is expected to alter the shape of the TW (Kanis and de Boer, 1994). Shera and Zweig's theory predicts that these changes should be accompanied by a change in spectral periodicity of SFOAEs.

Many authors have studied the effects of suppressor tones on OAEs (e.g., Kemp and Chum, 1980a; 1980b; Brass and Kemp, 1993; Zwicker and Wesel, 1990; Sutton, 1985; Dallmayr; 1987; Tavartkiladze *et al.*, 1994). However no experiments have looked for, or reported, a change in OAE spectral periodicity with suppressor level. It should be noted that with two tones the acoustic parameters define a four dimensional parameter space (stimulus frequency × stimulus level × suppressor frequency × suppressor level), but only experiments in a small subset of this space would be able to detect the predicted effect: those where stimulus frequency and suppressor frequency are swept together in a roughly constant ratio. Though this has been done (e.g., Kemp and Chum, 1980a) no change in

periodicity was reported. The effect of suppressor tones on SFOAE periodicity is a further effect that is investigated in this thesis.

2.17 Effect of Contralateral Noise, Ototoxic Drugs and Acoustic Overstimulation on OAEs

The ingestion of ototoxic drugs, the presentation of contralateral noise, and prolonged ipsilateral overstimulation are all thought to disturb the normal functioning of the active processes in some way, and therefore could, according to Shera and Zweig's theory, be accompanied by a change in spectral periodicity of SFOAEs.

Several authors have investigated the effect of ototoxic drugs on SFOAEs and TEOAEs (e.g., Long and Tubis, 1988; Karlsson *et al.*, 1991; Brown *et al.* 1993); of contralateral acoustic stimulation on OAEs (e.g., Collet *et al.*, 1994; Lind, 1994; and Giraud *et al.*, 1996); and of acoustic overstimulation on TEOAEs and SOAEs (Kemp, 1986). Whilst some clear changes in OAE characteristics have been induced, no changes in OAE periodicity have been reported.

3. Theories of OAEs

3.1 Cochlear Mechanics

The mechanical behaviour of the cochlea is still not fully understood. The development of theories of the cochlea has had several influences, including results from mathematical models of cochlear mechanics; results from direct mechanical measurements made in animal experiments; results from OAE experiments and results from psychoacoustic experiments. Direct measurements in the cochlea are difficult because the cochlea comprises complex and delicate structures embedded in bone, and because representative mechanical behaviour can only be measured in a living and undamaged cochlea. Also, the cochlear structures show a complex pattern of motion in three dimensions, whilst measurements of the motion are usually limited to only a few points. Consequently, measurements alone have so far been unable to completely elucidate the functioning of the cochlea. However, despite these difficulties, a consensus has been reached on the basic mechanisms of the cochlea. This is reviewed in Pickles (1988), Patuzzi (1996) and Dallos (1996), and a brief summary is presented here.

3.1.1 The Passive Travelling Wave

It has already been mentioned in section 1.1 that an acoustic stimulus in the ear canal generates TWs on the BM, which propagate along the cochlea from base to apex. The passive TW refers to the TW that results in those cases where the active processes are absent, or where their effects are insignificant, such as at high stimulus levels.

The passive structures of the BM, when uncoiled to lie along a linear axis, can be compared to a xylophone in that they behave as a series of beams, each with a different natural frequency. Although in the BM the "beams" are joined together to form a single membrane, the structural coupling between beams is thought to be very weak such that the beams can be treated as structurally independent. They are, however, mechanically coupled to each other by the motion of the cochlear fluid and it is this coupling that allows passive TWs to propagate along the BM. This passive TW wave propagation involves transfers of mechanical energy between a spatially distributed store of kinetic energy (in this case the cochlear fluids) and a spatially distributed store of potential energy (in this case the BM compliance). In this way it is physically similar to other mechanical wave motions such as surface water waves (Lighthill, 1981; Patuzzi, 1996).

The passive TW motion exhibits two interesting phenomena: frequency dispersion and spatial nonuniformity. Frequency dispersion, which results from the hydrodynamics of the system, means that the wave speed varies with frequency. Spatial nonuniformity means that the wave speed varies with position along the cochlea. This property, which arises from the variation along the cochlea of the structural properties of the BM forms the basis of passive frequency analysis performed by the cochlea. For any single frequency stimulus, there is a point along the cochlea where the natural frequency of the BM equals the stimulus frequency. As the TW approaches this point (known as the characteristic place for a given stimulus frequency), its speed of propagation falls off towards zero, and its amplitude rises to a peak. This leads to a peak in the mechanical (and consequently the neural) response in this region of the BM. In this way the cochlea maps stimulus frequency to position along the BM, with high frequency stimuli peaking near the base of the cochlea, and low frequency stimuli peaking near the apex. A more rigorous treatment of this behaviour based on the analysis by Lighthill (1981) and de Boer (1996), will be presented in part II.

Note that for a single stimulus frequency, the mechanics of the passive TW response involve both a wave motion along the BM and resonance phenomenon in a small region of the BM. This behaviour is far more complex than simple resonance at a point. Instead, the BM response arises from the propagation of a TW through a medium whose characteristic

impedance varies both with frequency and place, approaching zero at the point of resonance. In contrast to the simple second order filter obtained from a single degree of freedom resonator (which would arise if each point on the BM were coupled directly to the stapes), the TW response leads to a frequency response functions (at a given point) that has both a steeper high frequency cut-off and a much greater high frequency phase lag (3 to 5 cycles compared to 0.5 cycles). It was von Békésy's observation of this property of the cochlear response that led to the rejection of Helmholtz's resonance theory in favour of the travelling wave theory (Patuzzi, 1996).

3.1.2 The Cochlear Amplifier

It is currently believed that the propagation of the TW is enhanced by active processes involving the OHCs in the organ of Corti. The theory of an active mechanical amplification by some biological processes was first suggested by Gold[18](1948), though it was largely forgotten until being revived by Kemp (1978).

As was mentioned in section 1.1, current theory holds that the OHCs respond to an incoming TW by applying additional excitatory forces to the BM. Though this OHC motility is not well understood, it is believed that force generation arises from length changes in the OHCs. This system, known as the cochlear amplifier, greatly increases the TW amplitude thereby improving both the sensitivity and frequency selectivity of the ear. The cochlear amplifier can be viewed as creating a region of negative damping on the BM such that TWs propagating through this region (in either direction) acquire rather than dissipate energy. For a given stimulus frequency, this region is not thought to extend along the entire BM, but is instead located at and basal to the peak of the TW (de Boer, 1996). It follows that the location of this region is frequency dependent. As well as greatly improving the performance of the cochlea, the cochlear amplifier is also responsible for the generation of OAEs.

3.1.3 Nonlinearity in the Cochlear Amplifier

One important feature of the cochlea is that its response is strongly nonlinear. It is believed that the nonlinearity originates mainly in the cochlear amplifier, and is caused by

[18] Gold proposed active processes in order to explain the sharply tuned performance of the cochlea as a mechanical frequency analyzer. He also predicted the existence of SOAEs, though failed to measure these, probably because of equipment limitations.

saturation of certain processes in the OHCs (de Boer, 1996). Two manifestations of this are important in this project: the first will be termed self-suppression, and the second two-tone suppression. In self-suppression, a single pure tone is presented to the ear. As its level increases it is found that the BM vibration progressively approaches that of a fully passive (or dead) cochlea (e.g., Cooper and Rhode, 1992). (The passive cochlear response is thought to be approximately linear at moderate levels of stimulation.) In two-tone suppression (sometimes called 'mutual suppression') two tones of different frequency are presented simultaneously to the ear . The response of the BM at the frequency of the first tone is found to be reduced by the presence of the second tone (e.g., Rhode and Cooper, 1993) and vice-versa.

3.2 Development of Theories of SFOAE Generation

In the following sections, the development of the current theory of SFOAE generation due to Zweig and Shera (1995) is outlined. Theories of TEOAE and SOAE generation are also described, since, according to current theory, these are closely related to SFOAEs. Some alternative theories of SFOAE generation are also briefly discussed. DPOAEs are not covered in any depth as they are more complicated (Probst *et al.*, 1991; Shera and Guinan, 1999).

The discovery of OAEs and the demonstration of their cochlear origin (Kemp, 1978, 1979; Wilson, 1980a) raised several questions about cochlear mechanics. The presence of OAEs indicates both a source of energy within the cochlea, and a backward propagation of this energy. However, no such backward propagation of energy is seen in a large class of cochlear models when excited at the stapes. This remains the case even for models which include active processes and which show realistic BM mechanical responses (de Boer, 1996). A possible explanation for the failure of these cochlear models to predict OAEs arises from the reflection hypothesis, proposed by Kemp (1978), which is described in the next section.

3.3 The Reflection Hypothesis

In the reflection hypothesis, Kemp (1978) proposed that OAEs arise from an initially forward TW that is reflected at discontinuities in the BM mechanical impedance, thereby generating a backward TW which appears in the ear canal as an OAE. Kemp further proposed that the active processes play an essential role in OAE generation. At the time of its original suggestion, the theory was far from complete. For example, satisfactory

characterisations of both the cochlear amplifier and the putative discontinuities were still required. In this section, the main theories of OAE generation are reviewed. Most of these are directly descended from the reflection hypothesis in that they contain two essential elements: active processes and some form of reflection mechanism. Shera and Zweig's theory (1993b) and Strube's theory (1989) both fall into this category. The role of cochlear nonlinearity and multiple reflections of the TW in the cochlea are also discussed. Note that here the term 'reflection' covers more than simply passive reflection: it also covers the case where a forward TW stimulates an active source to generate or emit a backward TW.

It is first necessary to define the terms 'apical' and 'basal' TW reflection. A forward TW which encounters an impedance discontinuity on the BM will be reflected, giving rise to a backward TW. This process will be called apical reflection, and is essential for OAE generation. The backward TW then travels back along the BM until it reaches the stapes where it again encounters an impedance mismatch and is further reflected. This is called basal reflection and leads to multiple reflection in the cochlea (Kemp, 1980). Whilst basal reflection modifies the characteristics of the OAE signal, it is not essential to OAE generation. This is similar to the behaviour of the idealised ear shown in fig. 2.5. Apical reflection is discussed in this section, whilst basal reflection will be dealt with in section 3.6.

At the phenomenological level, the apical reflection can be characterised by the cochlear reflectance (looking apically), evaluated at the base of the cochlea, as in the idealisations in fig. 2.4 and 2.5 (Kemp, 1980; Shera and Zweig, 1993a). Recall from section 2.12 that the form of the cochlear reflectance has been estimated from OAE measurements. These indicate that the magnitude of the cochlear reflectance varies relatively slowly with frequency, whilst the phase angle drops rapidly, approximately following a $-\log(f)$ relationship. One aim of cochlear mechanical theories is to explain these characteristics.

One of the characteristics that received attention early on was the relatively long latency of OAEs (about 10 ms at 1 kHz), as compared to predictions of the latency based on estimates of TW travel time obtained from electrocochleographic measurements (e.g., 3.4 ms predicted from Eggermont's data, (1979)). Although they did not present a complete cochlear mechanical theory of OAEs, Kemp and Chum (1980b) argued that the long delays in the TEOAE could be explained by the steep sided mechanical filters that characterise the BM response at low stimulus intensities. As the stimulus level increases, the filter skirts become less steep giving shorter delays. This phenomenon, whereby latency reduces with increasing stimulus levels, has been used to explain the apparent

discrepancy between OAE and electrocochleographic measurements (e.g., Neely *et al.*, 1988).

Note that the OAE latency is sometimes regarded as having two components: the first associated with the BM resonant response near the characteristic place; and the second due to the TW transmission time to and from this place (e.g., Rutten, 1980). However, Strube (1989) points out that this split between TW propagation and BM resonance is artificial. Instead, the resonance of the BM is not considered to be a phenomenon with its own delay, but rather as one which influences the overall delay arising from TW propagation. Thus the speed of the TW varies continuously along the BM, and approaches zero in the region of BM resonance. This leads to long latencies for any TW originating within this region.

Originally Kemp (1978) had suggested that mechanical impedance discontinuities on the BM might be responsible for TW reflection. However, other possible mechanisms have also been explored. As part of the discussion of the mechanism of TW reflection, Kemp (1986) introduced the concept of 'place fixed' and 'wave fixed' source. A place fixed site is one whose location is independent of frequency, as would arise from a discontinuity in the BM impedance. A wave fixed source, perhaps arising from nonlinearity near the peak of the TW envelope, or from complex 2-D or 3-D fluid flows (Guelke and Bunn, 1985), would move with frequency as the TW peak moved. Kemp (1986) argued that the long group delays measured in TEOAEs and SFOAEs in humans are consistent with a place-fixed rather than a wave-fixed hypothesis.

This argument against a purely wave fixed generation mechanism has been restated by Strube (1989), Zweig and Shera (1995) and Shera and Guinan (1999), based on what is known as the scaling symmetry (or shift-symmetry) of the cochlea. Essentially this means that the TW spatial response pattern at any stimulus frequency resembles that at any other frequency (both in magnitude and phase), except for a translation in place along the BM. Thus, where scaling symmetry holds, knowledge of the BM response as a function of place at any single stimulus frequency is sufficient to define the BM response at all frequencies and all places. Alternatively, knowledge of the BM response as a function of stimulus frequency at any single place is also sufficient to define the BM response at all frequencies and all places. It has been shown that the group delays of any wave-fixed emissions from a scaling symmetric cochlea are very short, and therefore any ripple pattern will have a much wider peak-to-peak frequency interval (i.e., a greater periodicity) than that observed experimentally (Strube, 1989). This result holds true for any wave-fixed generation mechanism in an approximately scaling symmetric cochlea. This is discussed further in part II.

Assuming that a place-fixed mechanism is involved, the question remains as to what form of reflection site could explain the empirical results. Three forms are discussed below: isolated point reflection sites, spatially periodic reflection sites, and spatially random reflection sites.

Note that various different terms relating to the mechanism of TW reflection are found in the literature. The general terms "reflection site" and "scattering site" are used for to describe either wave-fixed or place fixed mechanisms. Both terms are used in this thesis. The terms "BM discontinuity", "irregularity" and "inhomogeneity" are used to describe a place fixed mechanism arising from the deviation of the BM impedance (or more strictly the cochlear wave impedance) from the smooth impedance variation associated with the mechanical frequency analysis performed by the cochlea. In this thesis, the term "BM inhomogeneity" is generally used. The terms "irregularity" and "discontinuity" would be misleading if applied to the smooth, periodic impedance variation proposed by Strube (1989), and will henceforth be avoided.

3.3.1 Point Reflection Sites

Zwicker and Lumer (1985) simulated OAEs from one, two and three localised inhomogeneities on the BM. A model with single inhomogeneity has two failings. First the predicted OAE spectra are only seen in a very limited frequency range. This is because it is only at these limited frequencies that the TW envelope is large in the region of the reflection site. Secondly, the predicted group delay is too short, except around one stimulus frequency (the characteristic frequency of the reflection site), where the peak of the TW coincides with the reflection site. At this frequency, the group speed of the TW approaches zero at the reflection site, hence increasing the predicted group delay. With three inhomogeneities, the predicted periodicities could be made to roughly agree with observed values, but the predicted OAE spectra still looked unnatural. Zwicker (1986, 1988, 1990) noted that the SFOAE periodicity (or equivalently the group delay of the TEOAEs) depended on the wavelength of the TW near the peak. Several other authors have published essentially similar results from various models with a few discrete inhomogeneities (e.g., Zwicker and Lumer, 1985; Zwicker and Peisl, 1990; Furst and Lapid, 1988; Fukazawa, 1992, Fukazawa and Tanaka, 1996; Wada *et al.*, 1999).

3.3.2 Spatially Periodic Inhomogeneities: Strube's Theory

Manley (1983) suggested that the periodicity seen in the auditory microstructure, in SFOAEs and in SOAEs could be explained if the strength of the active cochlear processes varied periodically along the length of BM. A somewhat similar theory was proposed by Strube (1989), following the work of Strube (1985), Zwicker (1986, 1988), and Peisl (1988). In this theory (referred to henceforth as "Strube's theory") it is postulated that some of the BM mechanical properties vary periodically with place (over and above the exponential variation associated with place-frequency mapping). This spatially periodic impedance variation is sometimes referred to as a BM corrugation, and gives rise to reflections (or scattering) of the forward TW. Strube compared the scattering mechanism to a phenomenon known as Bragg scattering, whereby a wave is strongly reflected when it encounters a sinusoidal spatial impedance variation whose spatial period equals half the wavelength (the Bragg condition).

To see how the Bragg condition is satisfied over a wide range of stimulus frequencies, consider a BM corrugation with a spatial period significantly less than half the TW wavelength at the base of the cochlea (for a large range of stimulus frequencies). For any given stimulus frequency, the wavelength of the TW decreases continuously as the TW propagates forward, and approaches zero as the TW approaches its characteristic place[19]. Consequently there is a point on the BM, basal to the characteristic place, where the Bragg condition is satisfied, and Bragg scattering occurs. If this point lies close to the TW peak then the amplitude of the backward scattered TW becomes significant. By scaling symmetry, the TW wavelength at the BM peak is independent of stimulus frequency (since changing the frequency simply shifts the response pattern along the BM, without changing its shape). Therefore the Bragg condition is satisfied for a point near the TW peak over a wide range of stimulus frequencies.

It should be noted that, although this Bragg point moves with the peak of the TW as the stimulus frequency is varied, the spatial pattern of BM inhomogeneities remains place fixed as is required in order for the predicted OAEs to have the long group delays observed in measurements. The mechanism by which the backward wave is generated is complex, involving the summation of reflected wavelets over the region of the TW peak, rather than just one single localised reflection. The spatial variation in the BM impedance ensures that all the reflected wavelets combine in phase, leading to a significant total reflection.

[19] The wavelength is only actually zero when the BM is in resonance. This is point lies slightly apical to the TW peak.

The theory successfully predicts the observed spectral periodicity, which is determined by the spatial periodicity in the BM properties. The theory does not rely on any nonlinearity, but does include active (though still linear) processes. However, these are only required to explain the magnitude of measured OAEs, rather than the essential reflection (or scattering) mechanism. Manley's suggestion (1983) that the corrugations be in the BM structures associated with the active processes is sufficient, but unnecessary in Strube's theory. Corrugations may equally well be in the passive structural properties of the BM.

In humans, the wavelength in the TW peak region is around 900 μm, and hence to satisfy the Bragg condition, Strube proposed a spatial period of around 500 μm for the BM impedance variation. However, in anatomical examinations of cochleae, no evidence for any spatial periodicity in the cochlea has been found (Zweig and Shera, 1995).

3.3.3 Random Inhomogeneities: Shera and Zweig's Theory

Shera and Zweig (1993b) have proposed an alternative version of the place-fixed theory in which the inhomogeneities in the BM mechanical properties have a random spatial distribution along the length BM, with length scales down to about 10 μm (the typical width of a single hair cell). The approximately regular periodicity seen in SFOAEs then arises from a mechanism described as 'spatial filtering'. As in Strube's theory, the essential mechanism in Shera and Zweig's theory is linear scattering. Active processes are required in order to produce what Zweig and Shera describe as a 'tall and broad' TW envelope (1995). A more complete description of the theory is given by Zweig and Shera (1995) and Talmadge *et al.* (1998). In this theory, the observed spectral periodicity is determined by the TW wavelength in the peak region, whereas in Strube's theory it is determined by the spatial period of the corrugations. Shera and Zweig's theory is discussed in greater detail in part II.

As with Strube's theory, the mechanism by which the backward wave is generated involves the summation of reflected wavelets over the region of the TW peak, rather than just one single localised reflection. The spectral periodicity predicted by Shera and Zweig's theory are similar to those predicted by Strube's corrugated BM. A simplified explanation for this results is that the random spatial variation can conceptually be decomposed into its spatial Fourier components. Each Fourier component appears as a corrugation with a different spatial period, and each one acts to scatter the forward TW. However, only one of these corrugations approximately satisfies the Bragg condition near the peak of the TW, and therefore generates a strong backward wave. Scattering from all

the other spatial frequency components is negligible. Shera and Zweig call this mechanism coherent reflection filtering (Shera and Zweig, 1993b; Zweig and Shera, 1995).

An important difference between Strube's theory and Shera and Zweig's theory is the way in which the spectral periodicity arises. In Strube's model (1989), it arises from the regular wave-like corrugation along the BM, which has a typical wavelength in humans of around 500 μm. In Shera and Zweig's model, it arises from an interaction between the TW and a random array of inhomogeneities along the BM. This means that the SFOAE periodicity predicted by Strube's model is fixed by the BM corrugations, whilst in Shera and Zweig's model it depends on the shape of the TW envelope.

Recall from section 2.15 that there is evidence that the speed of TW propagation increases with stimulus level. It might then be argued simplistically that an increase in stimulus level should therefore cause both a reduction in the OAE group delay and an increase in the periodicity. This simple argument is based on the assumption that the group delay depends on the TW propagation time to and from a single place-fixed reflection site. As will be discussed in part II, this argument is in general inadequate in explaining the predictions of cochlear models. This is because, according to both Strube's and Shera and Zweig's theory, OAEs result not just from a single place fixed-reflection, but from the sum of many place-fixed reflections. In fact, Shera and Zweig's theory predicts that the group delay is dependent on the TW wavelength in the peak region, rather than to the overall TW travel time. However, since wavelength is related to TW speed, and it is the TW speed in the peak region that dominates the overall TW travel time, predictions based on Shera and Zweig's theory agree closely with results arrived at by the simple single-reflection site argument. In contrast, models based on Strube's theory predict no change in group delay or periodicity when TW propagation speed changes.

No quantitative anatomical evidence either for or against this random array inhomogeneities exists (direct measurement of the BM mechanical properties being impossible). However Zweig and Shera (1995) note that anatomical studies have found a "generalised irregularity" and "cellular disorganisation" in the arrangement of OHCs in the apical turns of the primate cochlea.

3.4 The Role of Active Processes in the Reflection Hypothesis

The versions of the cochlear reflection hypothesis proposed by Strube, and Shera and Zweig rely on both active processes and reflection sites. A question that has been addressed in the literature is why both active processes and reflection sites are required to

generate OAEs, when it might be argued that either one alone could generate some form of backward TW, and therefore an OAE.

The first part of the question can be restated more fully as follows. Why, if the cochlear amplifier is thought to exert forces directly on the BM, and if the BM supports TW propagation equally in both directions, does the cochlear amplifier not directly generate an OAE whenever it is activated by a forward TW? The answer to this can be found in the theoretical modelling results reviewed by de Boer (1996). For any given stimulus frequency, the cochlear amplifier creates a region where the characteristic impedance, as seen by the TW, has a negative damping component. This region extends basally from a point near the peak of the TW envelope. TWs passing through this region (in either direction) are progressively amplified. This can be seen as a reversal of the more common process by which an acoustic wave decays as it propagates through a lossy medium. Such a wave is dissipated without being reflected. In the cochlea, the direction of energy flow is reversed to give a wave that blooms as it propagates, but is not reflected in the process. This is most easily illustrated mathematically by considering the cochlear amplifier to be linear. The presence of the cochlear amplifier can then be incorporated into a version of linear wave equation, where the negative dissipation manifests itself as a negative imaginary part to the wave impedance (which is related to the BM resistance). It can then be shown that, with an initial forward wave, backward waves only arise from rapid spatial changes in the wave impedance, and not from the sign of the resistance (de Boer, 1996, p. 274). The physical explanation corresponding to the mathematics is as follows. Although each infinitesimal element of the cochlear amplifier exerts a force on the BM which, in isolation, would generate travelling wavelets propagating in both directions, the spatial distribution and phase of these wavelets (which are both determined by the incident forward wave) are such that all the wavelets add together to give an additional forward wave which enhances the incident wave, but no backward wave.

The second part of the question is this. Why do backward TWs not arise from the scattering of the passive TW? (These backward TWs would not qualify as OAEs, according to the definition given in section 2.1, but they might be expected to produce a similar ripple pattern in the ear canal pressure to that observed). This question is addressed by Zweig and Shera (1995). They argue that in a purely passive cochlea, any TW reflections decay too rapidly to appear as significant backward waves arriving in the ear canal. Furthermore, simply reducing the level of damping assumed in the cochlea does not solve the problem, because the resulting TW envelope then appears to be too sharp. Instead it is argued that large backward TWs only arise when the TW envelope is both 'tall and broad', as it is when certain forms of cochlear amplifier characteristics are included.

'Tall' means that the forward TW is sufficiently strong to generate a significant backward wave. 'Broad' means that the TW peak region spans at least one full wavelength, such that Bragg scattering can arise. When the TW peak is too narrow, as it would be in a passive cochlea with very low damping, then the TW wavelength varies very rapidly over the peak region, because of the minimum phase property of the TW function discussed in part II. This means that there is no one dominant wavelength in the peak region, leading to a very broad band 'spatial filter'. This is inconsistent with the observed spectral periodicity. At first sight, this argument may appear to run counter to the widely accepted fact that active processes give rise to a sharp rather than a broad envelope. In fact with Zweig's cochlear amplifier (1991) the envelope is sharper than that arising in a passive cochlea with realistic damping, but is broader than that which would be seen in a cochlea which had very low damping, as could arise if the cochlear amplifier simply acted to provide active undamping. This is discussed in more detail in part II.

3.5 The Role of Cochlear Nonlinearity

The versions of the cochlear reflection hypothesis proposed by Strube, and Shera and Zweig do not rely on any cochlear nonlinearity for the generation of SFOAEs. The validity of this simplification is supported by the measurements of Shera and Zweig (1993a) which showed that the SFOAE spectrum varies linearly with stimulus level at very low stimulus levels. Thus, SFOAEs exist even when nonlinearity is absent.

At higher stimulus levels OAEs show a compressive nonlinearity, which is thought to arise from nonlinearities in the cochlear amplifier response. Results of click suppression and pure tone suppression experiments reported by Kemp and Chum (1980b), which indicated that the OAE response comprised a linear signal delay followed by a saturating nonlinear element, are consistent with this hypothesis.

Recently, Talmadge *et al.* (2000) have presented an analysis of Shera and Zweig's theory based on nonlinear cochlear models. This generally supports the argument that the long delays seen in OAEs arise from essentially linear reflection. However, they also report that a purely nonlinear component of SFOAEs can arise, even in perfectly smooth cochleae. This originates from deviations in the nonlinear cochlear mechanics from scaling symmetry and it manifests itself as a very short-latency (and therefore high periodicity) SFOAE component. Whilst at low and moderate stimulus levels this component is small, in some cases it can predominate at high stimulus levels. This is important for this project where the changes in periodicity that are of interest are those related to changes in the TW wavelength, rather than to additional nonlinear effects.

Thus in the theories under consideration here, nonlinearity is not essential for OAE generation, but is an important phenomenon at all but the lowest stimulus levels.

3.6 The Role of Multiple Reflections and the Middle-ear

Kemp (1980) presented an analysis relating evoked OAEs to the transmission characteristics of the middle ear, and to apical and basal reflections in the cochlea. This analysis predicted the presence of multiple reflections in the cochlea, whereby the forward wave is reflected at some site on the BM to generate a backward TW, which is in turn reflected at the stapes to generate a second forward wave and so on. This is similar to the situation illustrated in fig. 2.5.

At certain frequencies this second forward TW interferes constructively with the original forward TW, leading to a resonance[20] phenomenon whereby both forward and backward TWs are stronger at certain frequencies. This occurs when the product of the apical and the basal reflectances is purely real. Kemp (1980) and Wilson (1980a) suggested this as explanation of the peaks in the spectrum of TEOAEs, SFOAEs and the auditory microstructure. Kemp (1980) also pointed out that active processes may give rise to an apical reflection coefficient greater than unity (i.e., the backward TW may contain more energy than the original forward TW). If, at a particular frequency, the product of the apical and basal reflection coefficients is a real value exceeding unity then the cochlea will be unstable (i.e., any infinitesimal perturbation will grow unbounded at that frequency). Kemp (1980) proposed this as an explanation of SOAEs (where in practice the growth is limited by nonlinear saturation in the cochlear amplifier). This phenomenon can be explained by consideration of the two systems in figs. 2.3 and 2.5, which show multiple reflections. If the forward looking reflectance were increased in magnitude beyond unity, then a point would be reached where the train of reflection components seen in the echo (or OAE) response (panel e) would not decay away, but would become a continuous periodic signal. The figure also shows how the frequencies of this self sustaining signal would correspond to the spectral peaks seen in panel f. By analogy, it would be expected that the frequencies of SOAEs would correspond to the peaks seen the spectrum of TEOAEs and SFOAEs. This is indeed what is found experimentally (Zwicker and Schloth, 1984; Probst

[20] An acoustical system is said to be resonant when its acoustic input reactance is zero, and its acoustic input resistance is a minimum. Thus for a high impedance loudspeaker (as is common), the resonances of the driven system correspond to minima in the pressure response measured by the microphone, whilst the antiresonances correspond to maxima, assuming a constant voltage drive to the loudspeaker.

et al. 1986, 1991). The explanation of audiogram fine structure is more complicated, and is not dealt with here.

When multiple reflections are included in Shera and Zweig's theory, the spectral periodicities seen in the audiogram fine structure, in SFOAEs, in TEOAEs and in SOAEs are successfully explained (Talmadge *et al.*, 1998). Note that the role of multiple apical and basal reflections differs in these four phenomena. Multiple reflections are essential to the explanation of spectral periodicity in SOAEs and the auditory microstructure, which arise from the resonance phenomena of the TW suggested by Kemp (1980). However, multiple reflections are not essential for the explanation of the basic spectral periodicity seen in the real and imaginary parts of SFOAEs. This basic periodicity corresponds to the slope of the phase of the frequency spectrum of the SFOAEs and TEOAEs. Here, only a single apical reflection is required. Subsequent basal (and further apical) reflections, if significant, modify the measurements such that periodicity is seen in the magnitude of the SFOAE, not just in the real and imaginary parts (Kemp, 1980; Shera and Zweig, 1993a; Talmadge and Tubis, 1993; Talmadge *et al.*, 1998). This is illustrated to some extent by the difference in the two responses of the idealised ear shown in figs 2.4 and 2.5. It is discussed more fully in part II.

The basal reflection coefficient is determined by the mechanical properties of the system comprising the middle ear, the ear canal and the termination of the ear canal by the OAE probe. Several authors have investigated the effect on OAEs of altering these properties. For example, Zwicker (1990) found that altering the impedance of the OAE probe changed the frequencies of the peaks in the audiogram fine structure, SFOAE spectrum and TEOAE spectrum.

As well as influencing the basal reflectance, these mechanical properties also determine the transmission characteristics through the middle ear, which influences OAEs measured in the ear canal. OAE transmission is most efficient in the frequency region of 1-2 kHz, with increasing loss outside this region (Kemp and Chum, 1980a; Kemp, 1980; Shera and Zweig; 1993a). In ears with abnormal middle ear pressure, or with middle ear dysfunction, OAEs may be altered or abolished completely (Probst *et al.*,1991). In normal ears, measured OAEs have also been found to be altered by the stapedial muscle contraction (Probst *et al.*,1991).

3.7 Alternative Theories to the Reflection Hypothesis

A number of other mechanisms have been suggested for generating OAEs. Wilson (1980b) proposed the cellular swelling model in which volume changes in the hair cells

stimulate the fast wave (the longitudinal compression wave) in the cochlea, rather than the slower moving backward TW. However, this theory fails to predict the observed spectral periodicity. (In the Shera and Zweig model, active processes cause the BM to deflect at constant thickness, rather than to swell. BM deflection couples to the slow TW propagation rather than the fast compression wave.)

Guelke and Bunn (1985) suggested that complex 3-D fluid flow in the region of the TW peak can act as a generation site. However, this leads to a wave-fixed primary source which has the problems discussed by Strube (1989).

Van Hengel and Maat (1993) and van Hengel *et al.* (1996) have proposed a complex theory which relies on instability in the active processes, on the nonlinear phenomena of suppression and entrainment, and on multiple reflection at the stapes. Unlike Shera and Zweig's theory, this theory uses a completely active cochlea in which all points are unstable when stationary. The introduction of one single strong emitter leads to complex nonlinear coupling between points, which can give rise to spectral periodicity. However, there are as yet no published results from the model showing realistic SFOAEs.

3.8 Summary

In the various versions of the reflection hypothesis that have been proposed the involvement of the following mechanisms has been proposed: cochlear nonlinearity, BM inhomogeneities, 3-D fluid flows, active amplification, and multiple reflections within the cochlea. However, of the theories mentioned, only Strube's and Shera and Zweig's have been developed far enough for detailed models to be constructed which predict SFOAEs with realistic periodicities. These two theories are linear, active and 1-D but invoke different spatial distributions of BM inhomogeneities to explain SFOAE generation.

A key difference between Strube's theory and Shera and Zweig's theory is the way in which the spectral periodicity is determined. In Strube's model (1989) it is fixed by the spatial period of the BM corrugations, whilst in Shera and Zweig's model (1993b) it depends on the shape of the TW envelope, and therefore may be altered depending on the strength of the cochlear amplifier.

PART II: THEORETICAL PREDICTIONS

4. Development of the Cochlear Macromechanical Model

4.1 Overview

The objective in developing a cochlear model in this thesis is to illuminate the various roles played by the different elements in the theory, and where possible to relate these elements to measurable characteristics of SFOAEs. Moreover, the model should enable predictions based on the current theories of OAE generation to be made, which can then be tested experimentally. To achieve this objective, the model must capture all the essential features of the theory, whilst eliminating any non-essential features which would both confuse the interpretation of the results and increase the computational burden. The approach taken here has been to start with as simple a model as possible, and then to add complexity to check whether any significant changes result.

Cochlear mechanics is often divided into macro and micromechanics. Macromechanics is concerned with the large scale motions of both the cochlear fluids and the BM, whilst micromechanics is concerned with the detailed motions of the many components in the organ of Corti. The models considered in this thesis provide a mathematical representation of the macromechanical response of the cochlea only. Although cochlear micromechanical behaviour does have an influence on the macromechanical response, it need not be included explicitly. Instead, micromechanical effects can be characterised as relationships between macromechanical quantities. The justification for developing a macromechanical model is that, according to the underlying theory, the characteristics of measured SFOAEs are determined by the macromechanical behaviour of the cochlear responses consisting of TW propagation, amplification, dissipation and reflection. Thus, for example, the outer hair cells are not modelled directly but are instead included through the mechanical properties of the cochlear amplifier.

In the following section, details of the macromechanical cochlear model used in this thesis are presented and discussed. This is a 1-D, longwave model of the cochlea, including a locally active cochlear amplifier and BM inhomogeneities. A model variant including a frequency domain implementation of cochlear amplifier nonlinearity is also described. All the elements in the models are based on a features found in various published models. However, these element have not been brought together before as

described here. As well as specifying the model, in the following section relevant aspects of the model behaviour are also discussed.

4.2 Basic Assumptions in the Macromechanical Model

In the models used here the following simplifications have been made. The spiral shape of the cochlea is unrolled giving a straight cochlea. Reissner's membrane is ignored, and thus the scala vestibuli and scala media are treated as a single fluid channel, called here the upper channel. A section across the upper channel (perpendicular to the longitudinal axis) is assumed to be rectangular. The scala tympani, called here the lower channel, is assumed to have the same cross section as the upper channel. The basilar membrane, the tectorial membrane, the organ of Corti and associated support cells are all replaced by a single flexible membrane called the cochlear partition (CP). It is motion of the CP that displaces the cochlear fluids. The CP is assumed to be incompressible: it deflects, but it does not change volume. Thus the fluid velocities above and below the CP are equal at any location. The stapes footplate forms the basal boundary of the upper channel, and is perpendicular to the longitudinal axis. Structurally, there is no longitudinal coupling between points on the CP. Thus the CP is viewed as a series of independent beams somewhat like a xylophone, but with no gaps between the keys. The helicotrema is modelled as a gap in the CP at the apex of the cochlea. The cochlear fluids are assumed incompressible and inviscid. In this report, the terms 'cochlear partition', 'upper channel' and 'lower channel' are adopted for the model representations. In the literature, these are sometimes referred to as the 'basilar membrane', 'scala vestibuli' and 'scala tympani' respectively. These simplifications are common in many discussions of cochlear mechanics (e.g., de Boer, 1996; Patuzzi 1996). The validity of these simplifications is discussed below.

4.2.1 The Assumption of Incompressible Flow

The assumption that the fluid is incompressible is justified provided that the stimulus frequency is low enough (Patuzzi, 1996). For a stimulus frequency of 3 kHz, and taking the speed of sound in the cochlear fluid as 1500 m/s, and the length of the fluid channels as 35 mm, the resulting longitudinal wave has a wavelength of about 500 mm or about 14 times the length of the cochlea, and therefore can be safely ignored.

4.2.2 The Assumption of Uniform Scalae

As is usual in cochlear models (e.g., Neely and Kim, 1986; Kanis and de Boer, 1994; Talmadge *et al.* 1998), the cross-sections of the scalae are assumed to be independent of the axial location. Shera and Zweig (1991a) argue against this assumption, pointing out that, if lower frequency behaviour is to be accurately represented, the channel cross-section should taper along the longitudinal axis of the cochlear model. The effect of this taper is to maintain a resistive cochlear input impedance at low frequencies (and therefore good middle ear efficiency) and also to maintain scaling symmetry in the cochlea (see section 4.7.4). However, the effect is only pronounced below about 1 kHz and is not considered here.

4.2.3 The Compressibility of the Cochlea

In this thesis the commonly made assumption that the volume velocity of the round window is equal and opposite to the volume velocity of the oval window at all frequencies of interest. This not only rules out significant fluid compressibility but also net fluid inflow or outflow from the cochlea via the aqueducts or blood vessels. An empirical check of this assumption is reported by Shera and Zweig (1992).

4.2.4 The Assumption of No Longitudinal Structural Coupling

Although there must also be some longitudinal structural coupling, this is ignored in the chosen models. Instead, points on the CP are only coupled to their neighbours via the fluid flow. For passive models, Lighthill (1981) argued that any longitudinal coupling must be very weak for the cochlea to function correctly. However, some active models do include a degree of longitudinal coupling associated with the active processes (e.g., Zwicker, 1988; Fukazawa and Tanaka, 1996). However, as with much of cochlear modelling, there appears to be little direct empirical evidence either for or against these assumptions. In this thesis, following most cochlear modellers, all longitudinal coupling has been ignored as an unnecessary complication (e.g., Kanis and de Boer, 1993b; Zweig 1991).

4.3 Defining the Macromechanical Variables

Fig 4.1 shows a cross section through the cochlea. Here the passive structures of the CP are represented by a simple spring-mass system, and the cochlear amplifier is represented by a motion sensor, a filter, and a force generator. The filter may be linear or nonlinear. The basic place-frequency mapping in the cochlear model arises from the (roughly) exponential reduction in local natural frequency of the CP along the cochlear axis. This may be achieved either by reducing the CP local stiffness at constant CP mass, or by both reducing the CP local stiffness and increasing the CP local mass along the cochlea. (Here, 'local' mass and stiffness means mass and stiffness per unit length at a point on the cochlear axis.)

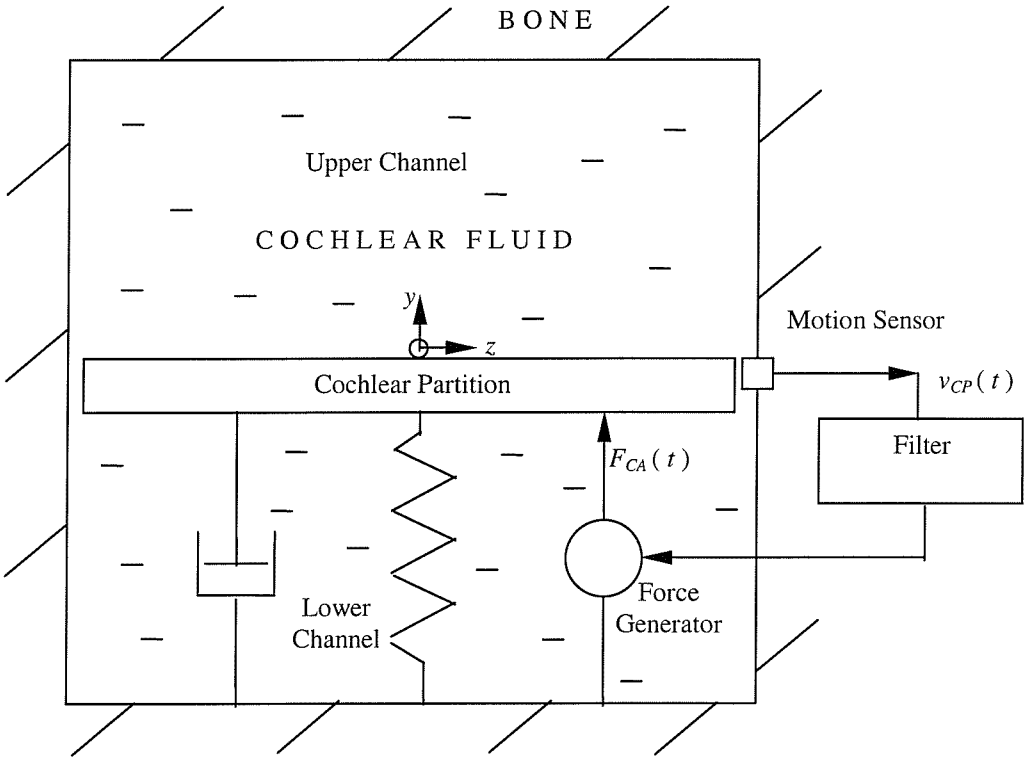


Fig 4.1. Cross section through the model of the cochlea, showing the mass-spring-damper representation of the cochlear partition. The cochlear amplifier is represented by a motion sensor, a filter and force generator. The sensor measures the CP velocity, $v_{CP}(t)$, leading to the generation of a force, $F_{CA}(t)$ applied to the CP.

For a given sound stimulus, the solution of the macromechanical response is specified in terms of the fluid pressure and fluid velocity vector at all points and all instants in time. Together, these define the flow field. Once the flow field is known, the CP velocity and pressure are also known, because the CP is in contact with the fluid.

Variables are specified with reference to the three-dimensional rectangular co-ordinate system shown in fig. 4.2 . The x-co-ordinate defines the longitudinal distance from the stapes; the y-co-ordinate the vertical distance from the CP; and the z-co-ordinate the lateral distance from the centreline of the CP. No lateral variation is considered in this 2-D analysis. At a given point in space and an instant in time, t, the flow field is denoted by:

the fluid pressure:	$p(x,y,t)$
the fluid x-velocity vector:	$u(x,y,t)$
the fluid y-velocity vector:	$v(x,y,t)$

The velocities at the fluid boundaries are denoted by:

the stapes x-velocity:	$u_{ST}(y,t)$
the round window x-velocity:	$u_{RW}(y,t)$
cochlear partition z-velocity:	$v_{CP}(x,t)$

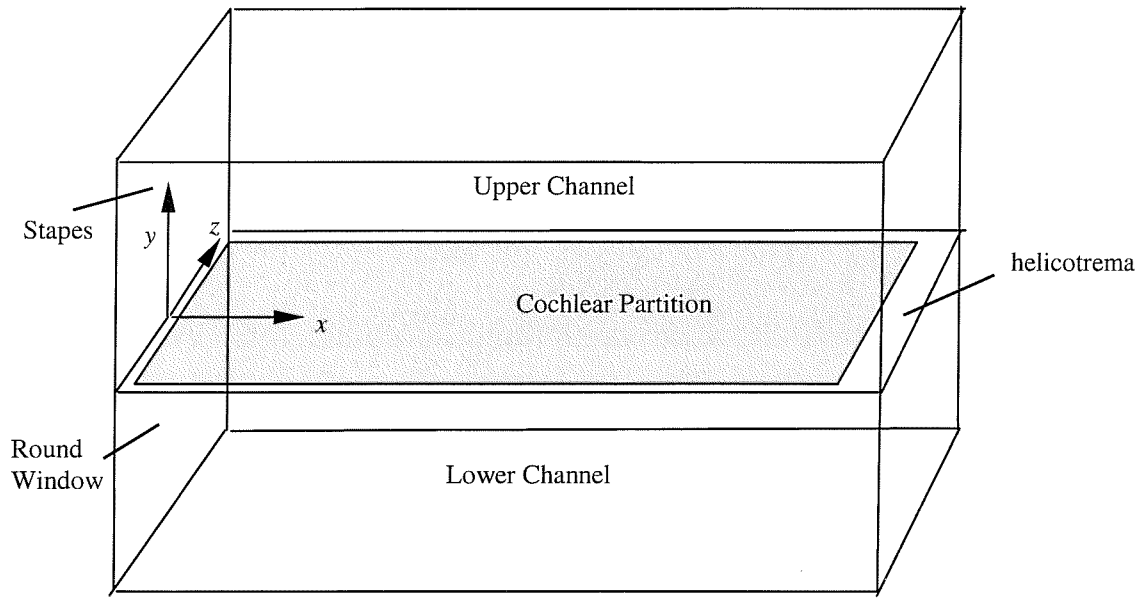


Fig 4.2. Definition of the co-ordinate system for the cochlear model.

4.4 Symmetry Relations in the Upper and Lower Channels

The symmetry of the upper and lower channels greatly simplifies the analysis. Patuzzi (1996) illustrates these relationships by imagining a cochlea with a second stapes in place of its round window, thus giving perfect symmetry. Applying identical forces to both stapes (the so called 'push-push' loading condition) causes compression waves (or 'fast' waves) to travel up the two cochlear channels. By symmetry, there is no associated CP displacement and therefore no cochlear TW. In contrast, applying equal and opposite forces to the two stapes (the 'push-pull' loading condition) gives rise to a pressure difference in the two channels thereby generating the familiar TW which propagates by CP deflection. It also generates two antisymmetric fast waves which travel up the two cochlear channels (for example, pushing the upper stapes and pulling the lower stapes generates a compression wave in the upper channel, and a rarefaction wave in the lower channel). In the real (i.e., single stapes) cochlea, any arbitrary boundary condition at the round window and stapes can be split into the sum of a push-push and a push-pull case (assuming linearity).

A further simplification arises when fluid compressibility can be ignored (i.e., at low stimulus frequencies). Here the fast waves can be ignored, since they propagate and decay away within a fraction of one stimulus period. In this case, push-push loading causes an instantaneous change in the pressure throughout the entire cochlear fluid (i.e., the pressure field remains entirely spatially uniform at all times). In the push-pull case, there is no overall change in the enclosed volume, and thus the fast wave can be ignored entirely. Thus in this case only the TW on the CP need be considered. Notice that in the push-pull case the pressure and velocity at apical locations remain completely unchanged until the arrival of the CP travelling wave: there is no energy transfer other than by the TW.

Mathematically this can be represented as follows. The pressures in the upper and lower channels can be combined into a semi-sum pressure, denoted p_s, and a semi-difference pressure, denoted p_d, which are then functions of the push-push and push-pull loading respectively:

$$\begin{aligned} p_s(x,y,t) &\equiv \tfrac{1}{2}p(x,y,t)+\tfrac{1}{2}p(x,-y,t) \\ p_d(x,y,t) &\equiv \tfrac{1}{2}p(x,y,t)-\tfrac{1}{2}p(x,-y,t) \end{aligned} \quad \text{for } 0<y<H \qquad [4.1]$$

This definition is convenient because (assuming fluid incompressibility) only the semi-difference pressure, p_d, plays a part in TW propagation. The semi-sum pressure, p_s, merely acts as an additional spatially uniform term. For completeness the value of p_s will be derived in this section, but thereafter attention is turned exclusively to p_d.

The semi-sum pressure depends on the push-push loading component. In the real cochlea, this component has a particularly simple form if, as is common, it is assumed that the impedance at the round window is approximately zero. From this, it follows that, at any instant, the uniform semi-sum pressure is simply half the instantaneous pressure at the stapes, as is shown below:

$$
\begin{aligned}
p_{RW}(y,t) &\equiv p(x,y,t)\,\big|_{x=0;\,y<H} = 0 \quad \text{for all } t \\
p_{St}(y,t) &\equiv p(x,y,t)\,\big|_{x=0;\,y>H} \\
p_s(t) &= \tfrac{1}{2}p(x,y,t) + \tfrac{1}{2}p(x,-y,t) = \tfrac{1}{2}p_{RW}(t) + \tfrac{1}{2}p_{St}(t) \\
&= \tfrac{1}{2}p_{St}(t)
\end{aligned}
\qquad [4.2]
$$

where p_{RW} and p_{St} are the pressures at the round window and stapes respectively. This is shown more rigorously in appendix I.

Fluid incompressibility also means that the fluid velocity is completely unaffected by the push-push case. The velocity components then exhibit the following antisymmetry:

$$
\begin{aligned}
u_{RW}(-y,t) &= -u_{ST}(y,t) \\
u(x,-y,t) &= -u(x,y,t) \\
v(x,-y,t) &= v(x,y,t)
\end{aligned}
\qquad [4.3]
$$

These symmetry relations are shown in fig. 4.3. This time domain representation leads to an analogous symmetry relation in the frequency domain for the real and imaginary parts of the Fourier transforms of the velocity components.

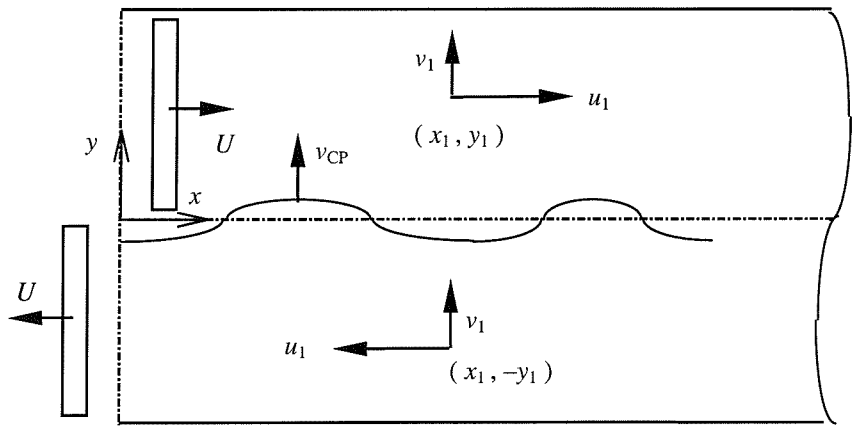


Fig 4.3. Symmetry relations in the cochlear model for instantaneous velocities. Arrows indicate actual physical direction of motion. The quantities x_1, y_1 are positive displacements, whilst U, u_1 and v_1 are positive velocity scalars. The quantities u (x , y) , v (x, y) , u_{St} and u_{RW} are signed velocity components, defined as positive in the direction of the x-y co-ordinate system. All velocities are functions of time.

4.5 The Longwave Model and the 1-D Wave-Equation

In this section, the equations governing TW propagation are introduced. Given the simplifying assumptions of the previous sections, the semi-difference pressure in the cochlear fluid can now be related to the displacement of the stapes via the equations of fluid mechanics (i.e., mass and momentum conservation) and the impedance of the CP. Fluid viscosity will be ignored (though the system remains damped via the resistive component of the CP impedance). The equations are linearized, since the particle displacements and velocities are very small. The full derivation is presented by de Boer (1996).

The analysis yields partial differential equations for the fluid flow field in which both the axial and transverse spatial dimensions appear as independent variables. However there is a certain commonly used simplification, known as the longwave approximation, which eliminates the transverse co-ordinate as an independent variable. The mechanics of the resulting longwave model can then be represented as a simple one-dimensional wave equation. The full derivation of this wave equation is given in appendix I and can be found in de Boer (1996). Equation [4.4] gives the frequency domain version of the wave equation (i.e., the Helmholtz equation) for the semi-difference pressure, $\boldsymbol{p_d}(x, \omega)$, which, in the longwave model, is independent of y:

$$\frac{d^2 \boldsymbol{p_d}}{dx^2} + \boldsymbol{k}^2 \boldsymbol{p_d} = 0$$

$$\boldsymbol{k} \equiv \frac{\omega}{\boldsymbol{c_{TW}}} \qquad [4.4]$$

$$\boldsymbol{c_{TW}^2} = \frac{i\omega H \boldsymbol{Z_{CP}}}{2\rho}$$

where, $\boldsymbol{p_d}$ is the Fourier transform of the semi-difference pressure across the CP, x is the distance along the CP, $\boldsymbol{k}$ is the wavenumber, $\boldsymbol{c_{TW}}$ is the TW phase speed, ω is the radian stimulus frequency, ρ is the cochlear fluid density, H is the height of the upper channel (here assumed constant), and $\boldsymbol{Z_{CP}}$ is the CP impedance (commonly known as the point mechanical impedance, though strictly it is the specific acoustical impedance). Note that $\boldsymbol{Z_{CP}}$, $\boldsymbol{k}$ and $\boldsymbol{p_d}$ are complex numbers (which will be denoted here by the bold typeface) and can be functions of both x and ω.

The semi-difference pressure must also satisfy two boundary conditions. For example, if the stapes velocity, $\boldsymbol{u_{St}}$, has been specified then the two boundary conditions are:

$$\left.\frac{dp_d}{dx}\right|_{x=0} = -i2\omega\rho u_{St} \quad \text{at the stapes} \qquad [4.5]$$

$$p_d|_{x=L} = 0 \qquad \text{at the helicotrema} \qquad [4.6]$$

The stapes boundary condition arises from the x-momentum equation which relates the pressure gradient to the fluid acceleration. It has also been assumed that the round window impedance is zero, which means that the true pressure at the stapes equals twice the semi-difference pressure at the stapes (equation [4.2]). As an alternative to [4.5], the stimulus could be specified in the ear canal, rather than at the stapes, if a middle ear model is also included (section 4.6.5).

The condition [4.6] at the helicotrema, though not very realistic, is nevertheless commonly used. The justification for this is that the model response is highly insensitive to this apical boundary condition, provided the stimulus frequency is high enough to ensure that the characteristic place lies basal to the helicotrema. This is because the TW decays almost entirely as it reaches its characteristic place, and thus very little energy reaches the helicotrema.

Equation [4.4], together with the boundary conditions [4.5] and [4.6] can be solved using a finite difference method to give the semi-difference pressure, p_d, arising from any specified stapes velocity. The CP velocity v_{CP} is also dependent only on x and ω, and can be obtained from the CP impedance once the semi-difference pressure is known:

$$v_{CP} = -2\frac{p_d}{Z_{CP}} \quad \text{(impedance relation for the CP)} \qquad [4.7]$$

Note that the factor of –2 in equation [4.7] arises from the sign convention and the definition of the semi-difference pressure in equation [4.1]

It is useful to split the CP point mechanical impedance, Z_{CP}, into three components:

$$Z_{CP}(x,\omega) = Z_{Pass}(x,\omega) + Z_{CA}(x,\omega) + Z_{Sc}(x,\omega) \qquad [4.8]$$

where Z_{Pass}, Z_{CA}, and Z_{Sc} are the components of impedance due to the passive cochlear structures, the cochlea amplifier, and any inhomogeneities (or scattering sites) respectively. The mechanics of the CP have here been assumed linear (to allow an impedance to be used).

4.6 Specification of the Model

The cochlear models are based on equations [4.4], [4.5], [4.6], [4.7] and [4.8] with the parameters specified below.

4.6.1 The Passive Cochlea

The four cochlear quantities required to solve equation [4.4] are specified here, assuming a passive CP impedance. These quantities are: H, L, ρ and $\mathbf{Z}_{Pass}(x,\omega)$. In this thesis, the values have been taken from two publications which detail longwave models of the human cochlea: Kanis and de Boer (1994) and Talmadge *et al.* (1998).

The first three quantities, which define the cochlear scalae dimensions and the cochlear fluid, are simply specified by three parameters. Models based on Kanis and de Boer (1994) use H = 1 mm whilst those based Talmadge *et al.* (1998) use H = 3.79 mm. Both models assume ρ = 1000 kgm^{-3}. The value L = 35 mm has also been taken. This is unimportant, provided it is significantly greater than the highest characteristic place of interest.

The fourth quantity is the passive CP impedance, $\mathbf{Z}_{Pass}(x,\omega)$, which characterises the structural properties of a point on the CP in isolation both from the rest of CP and from the cochlear fluids. The passive structure at each point is modelled by a simple mass-spring-damper system, whose point impedance is then specified by three parameters: mass, stiffness and damping rate. Alternatively, the point impedance can be specified in terms of mass, natural frequency and critical damping ratio, as in [4.9]:

$$\mathbf{Z}_{Pass}(x,\omega) = \frac{m_0}{i\omega}\omega_C^2(x) + m_0\delta(x)\omega_C(x) + i\omega m_0 \qquad [4.9]$$

where m_0 is the CP mass; ω is the stimulus frequency; ω_C is the natural frequency; δ is twice the critical damping ratio and x is the CP location. Note that δ is approximately equal to the reciprocal of the Q-factor[21] of the resonance peak. The passive CP impedance varies spatially such that its natural frequency, ω_C, varies (approximately) exponentially with position, rather like a xylophone. This gives rise to the basic place frequency mapping in the cochlea. The natural frequency is equal to the frequency at which the velocity amplitude response is a maximum for a given pressure amplitude. The natural frequency will be referred to here as the characteristic frequency[22] of a given point. The

[21] The Q-factor of the system is defined from the frequency response function of displacement per unit pressure. It is defined as the ratio of the frequency of the peak response to the 3 dB bandwidth. For $\delta << 1$, as here, it can be shown that Q-factor $\approx 1/\delta$ (Thomson, 1981).

[22] This definition of characteristic frequency does not strictly give the model equivalent of the definition used in cochlear physiological measurements, where it is defined for a given point as the frequency of maximal response (such as the velocity response) for a given excitation at some reference point such as the stapes. However, the two values will be close.

characteristic place is the inverse of this function: i.e., for a given pure tone stimulus, it is the place whose natural frequency equals the stimulus frequency.

The Kanis and de Boer model assumes a perfect exponential spatial variation of characteristic frequency, and a constant damping ratio:

$$\omega_C(x) = \omega_{C_0} \exp(-x/d)$$
$$\delta(x) = \delta_0 \qquad [4.10]$$

where d is the place-frequency length parameter. The parameters are given as: $d = 6.67$ mm; $\omega_{C_0} = 2\pi \times 22.508$ kHz; $m_0 = 0.5$ kgm^{-2}; $\delta_0 = 0.4$. This passive model will be referred to as the KdB-1994 model.

Talmadge *et al.* (1998) have a more realistic form of the passive impedance that includes a deviation from the perfect exponential characteristic impedance variation, and a spatially varying Q-factor:

$$\omega_C(x) = \omega_{C_1} \exp(-x/d) + \omega_{C_2}$$
$$\delta(x) = \frac{\delta_0 + \delta_1 \exp(x/d)}{1 + \alpha \exp(x/d)} \qquad [4.11]$$

where $d = 7.24$ mm; $\omega_{C_1} = 2\pi \times 20.8$ kHz; $\omega_{C_2} = -2\pi \times 0.1455$ kHz; $m_0 = 0.05$ kgm^{-2}; $\delta_0 = 0.0385$; $\delta_1 = 0.000765$; $\alpha = -0.007$. This passive model will be referred to as the T-1998 model.

The seemingly large differences between the parameters H and m_0 used in the two models can partly be explained by the assumed width of the CP. Note also that T-1998 is much more lightly damped than KdB-1994.

4.6.2 The Linear Cochlear Amplifier Impedance, Z_{CA}

The cochlear amplifier is characterised by its impedance, $\mathbf{Z}_{CA}(x, \omega)$, as a function of both place and stimulus frequency. It is generally believed that active processes in the organ of Corti create a region of negative damping near the peak of the TW. This corresponds to a $\mathbf{Z}_{CP}$ with a negative real part in this region of the CP (for a given stimulus frequency). This in turn leads to a wavenumber with a negative dissipative component (via equation [4.4]), which amplifies propagating waves in this region. The argument for this form of the active impedance is discussed in detail by Zweig (1991) and de Boer (1996), and is based on attempts to fit results from cochlear models to *in vivo* measurements of the mechanical response of the mammalian cochlea.

Two different forms of the cochlear amplifier impedance have been adopted. The first is based on the model of Neely and Kim (1986), but with the precise formulation and parameters taken from Kanis and de Boer (1994). The second is based on Zweig's model (1990, 1991) with the parameters taken from Talmadge *et al.* (1998).

The (linear) cochlear amplifier impedance reported in Kanis and de Boer (1994) is:

$$Z_{CA}(x,\omega) = e_0 d_0 \omega_C(x) \frac{1 + i\beta(x,\omega)}{\delta_{SC} + i\left[\beta(x,\omega) - \sigma^2 / \beta(x,\omega)\right]}$$
$$\beta(x,\omega) \equiv \frac{\omega}{\omega_C(x)} \qquad [4.12]$$

where $e_0 = 4.28\times10^{-5}$ kgm^{-2}s; $d_0 = 1404$ kgs^{-1}; $\delta_{SC} = 0.14$; $\sigma = 0.7$; and ω_C is as given in [4.10] for the KdB-1994. This impedance will be referred to as the Neely and Kim cochlear amplifier, abbreviated to NK-1986, in this thesis.

The (linear) cochlear amplifier impedance formulated by Zweig (1990, 1991) and reported in Talmadge *et al.* (1998) is defined by:

$$Z_{CA}(x,\omega) = \frac{m_0 \rho_f \omega_C^2(x)}{i\omega} \exp\left[-i\psi_f \beta(x,\omega)\right] + \frac{m_0 \rho_s \omega_C^2(x)}{i\omega} \exp\left[-i\psi_s \beta(x,\omega)\right]$$
$$\beta(x,\omega) \equiv \frac{\omega}{\omega_C(x)} \qquad [4.13]$$

where m_0 and ω_C are defined as for the passive model [4.9], [4.10]. The parameter values are taken from Talmadge *et al.* (1998): $\rho_f = 0.16$; $\rho_s = 0.1416$; $\psi_f = 2\pi\times0.24$; $\psi_s = 2\pi\times1.742$ (the subscripts '*s*' and '*f*' referring to slow and fast feedback terms). This impedance will be referred to as the Zweig cochlear amplifier, abbreviated to Z-1991.

4.6.3 The Scattering Impedance, Z_{Sc}

Three spatial forms of scattering impedance have also been modelled: the spatially random inhomogeneities proposed by Shera and Zweig (1993b); the spatially periodic corrugations proposed by Strube (1989); and (for comparison) a single CP impedance discontinuity to generate a point reflection. In addition, a variant of Strube's model has also been used, where the corrugations are amplitude modulated in space, to give a narrow-band rather than a purely sinusoidal variation.

The scattering impedance has been generated by starting with the passive damping rate as a function of CP place, and multiplying this by a spatially varying perturbation function. For Shera and Zweig's model, the perturbing function is a gaussian random

'spatial white noise' sequence; for the Strube model it is a sinusoidal function of place; and for a point reflection site, the perturbation is a step function.

The use of the damping rate ensures that the scattering impedance is independent of stimulus frequency, which means that the simplified phenomenological model due to Zweig and Shera (1995) (to be discussed in section 4.8) becomes more applicable. The effect of using frequency dependent forms of the scattering impedance based on perturbing combinations of the CP mass, stiffness, or cochlear amplifier impedance have also been explored. One of these alternative cases, in which the overall CP impedance is perturbed, is reported here.

Mathematically the scattering impedance is defined in the following equations. In cases [4.15] to [4.18], it is obtained by perturbing the passive damping term at any point, $r(x)$, which is determined by the mass, natural frequency and damping ratio:

$$r(x) = m_0 \delta(x) \omega_C(x) \qquad [4.14]$$

Five variants of the scattering impedance are detailed below.

(i) *The Shera-Zweig random scattering impedance*

$$\mathbf{Z}_{Sc}(x) = r(x) a_{Sc} b(x) \qquad [4.15]$$

where a_{Sc} is the scattering amplitude parameter typically set between 0.001-0.01. The spatial signal, $b(x)$, is a very broad band random gaussian signal with a RMS amplitude of 1, a lower cut-off spatial frequency given by $1/L$, and an upper cut-off spatial frequency given by 1/10 μm.

(ii) *The Strube spatially periodic scattering impedance*

$$\mathbf{Z}_{Sc}(x) = r(x) a_{Sc} \sin(2\pi x / l_{Sc}) \qquad [4.16]$$

where a_{Sc} is the scattering amplitude parameter typically set at about 0.005 and l_{Sc} is the scattering length (or spatial period), set to between 0.35 and 0.5 mm.

(iii) *The Strube narrow band scattering impedance*

$$\mathbf{Z}_{Sc}(x) = r(x) a_{Sc} n(x) \qquad [4.17]$$

where a_{Sc} is the scattering amplitude parameter typically set between 0.001 and 0.01. The spatial signal, $n(x)$, is a narrow band random gaussian signal with an RMS amplitude of 1, a centre spatial period of 0.5 mm, and a spatial bandwidth of 2 mm^{-1}.

(iv) *The single point reflection site*

This is achieved with a step function in the impedance

$$\mathbf{Z}_{Sc}(x) = r(x) a_{Sc} \operatorname{sgn}(x - x_{Sc}) \qquad [4.18]$$

where a_{Sc} is the scattering amplitude parameter typically set to about 0.03. x_{Sc} is the location of the discontinuity and "sgn" is the signum function, which has a magnitude of 1 and a sign equal to the sign of its argument.

(v) *The Shera-Zweig frequency dependent random scattering impedance*

In order to investigate the influence of the frequency dependence of the scattering impedance, an alternative form of the random scattering has also been used, based on perturbations of the entire CP impedance rather than the damping:

$$\mathbf{Z}_{Sc}(x,\omega) = \left[\mathbf{Z}_{Pass}(x,\omega) + \mathbf{Z}_{CA}(x,\omega)\right] a_{Sc} b(x) \qquad [4.19]$$

where $\mathbf{Z}_{Pass}(x,\omega)$ and $\mathbf{Z}_{CA}(x,\omega)$ are the CP passive impedance and cochlear amplifier impedances, defined in [4.8]; a_{Sc} is the scattering amplitude parameter typically set between 0.001-0.01; and $b(x)$ is a very broad band random gaussian signal with an RMS amplitude of 1, a lower cut-off spatial frequency given by $1/L$, and an upper cut-off spatial frequency given by 1/10 μm.

4.6.4 Nonlinearity in Cochlear Models

Zweig and Shera's analysis of SFOAE generation (1995) is based on a linear cochlear model. However, to test the predictions of this analysis, this thesis uses the nonlinear phenomena of self-suppression and two-tone suppression to modify the shape of the TW. In order to predict the SFOAE frequency functions in these two nonlinear cases, the quasilinear model of Kanis and de Boer (1993b, 1994, 1996) was modified to include CP inhomogeneities. This model includes a representation of compressive nonlinearity in the OHCs, but treats the nonlinearity using a simplified, 'quasilinear' method. It is capable of predicting self-suppression, two-tone suppression, and distortion product OAE generation.

The quasilinear method works in the frequency domain and is numerically very efficient when only a few (in our case two) primary tones are present. It works iteratively as follows. First the unsuppressed velocity response due to one tone is calculated as for the linear model. This includes calculating the active pressure due to the cochlear amplifier at each point on the CP, assuming no saturation. This is the notional active pressure, denoted $\mathbf{p}_{CA,Lin}$, that would arise for the given CP velocity in the absence of saturation. This is then converted into a waveform in the time-domain, and passed though a saturating nonlinearity. Here, following Kanis and de Boer, saturating nonlinearity is modelled as a hyperbolic tangent function. The output of the saturating function is then converted back into the frequency domain, retaining only the primary frequency component

and discarding any distortion terms. This then defines a new value for the active pressure, denoted $p_{CA,QL}$ (for quasilinear) at each point. An effective quasilinear cochlear amplifier impedance, $Z_{CA,QL}$ can then be calculated at each point from $p_{CA,QL}$ and the CP velocity. Since saturation of the cochlear amplifier at any point affects the overall CP response, and hence feeds back on itself, the CP response must then be recalculated, with this new $Z_{CA,QL}$ in place of the linear value assumed initially. This process is repeated iteratively until the response stabilises. This then gives the self-suppressed cochlear response due to the first tone. This must then be repeated for the second tone (initially ignoring the presence of the first tone).

Having obtained the self suppressed responses due to the two primaries in isolation, their mutual suppression must be calculated. In this case, the two primary responses at each point are again converted back into the time domain and added together to give the total waveform of the CP active pressure. As before, this is passed though the hyperbolic tangent function to obtain the waveform of the quasilinear active pressure, which is then transformed to the frequency domain as before. The new effective $Z_{CA,QL}$ at the two frequencies is then calculated in the frequency domain, and used to recalculate the total CP response. This continues iteratively until the solutions for the two frequencies are stable. The quasilinear assumption remains valid provided that the amplitude of the CP response due to the presence of distortion products OAEs is much less than that due to the primary tones.

It should be borne in mind that the details of cochlear nonlinearity are not at all well understood, and that, as with all models, the model of Kanis and de Boer includes many assumed characteristics and parameter values. Therefore its predictions are unlikely to be quantitatively accurate, and should be treated with some caution. Despite this, this model may still produce informative results which are unobtainable with linear models. The nonlinear model was used to predict the modified shape of the TW in the self-suppression and two-tone suppression cases.

4.6.5 The Middle Ear, Ear Canal and OAE Probe Model

A model of the middle ear, ear canal and OAE probe are required to complete the full mathematical model, and allow SFOAEs in the ear canal to be predicted. For brevity, these three models will be referred to collectively as the 'front end model'.

The middle ear and ear canal can be represented by a two-port network (as in Kemp, 1980; Shera and Zweig, 1993a):

$$\begin{bmatrix} p_{St} \\ Q_{St} \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \begin{bmatrix} p_{EC} \\ Q_{EC} \end{bmatrix} \qquad [4.20]$$

where p denotes the acoustic pressure, Q the acoustic volume velocity, subscripts *EC* and *St* indicate locations at the entrance to the ear canal and stapes respectively, and the matrix elements T_{ij} define the transmission through the outer and middle ear. All terms are complex, and functions of frequency.

The OAE probe is represented as an acoustic volume velocity source, Q_{Src}, with an acoustic source admittance, Y_{Src}. The actual ear canal pressure and volume velocity are then given by:

$$p_{EC} = \frac{Q_{Src}}{Y_{Src} + Y_{EC}}$$
$$Q_{EC} = \frac{Q_{Src} Y_{EC}}{Y_{Src} + Y_{EC}} \qquad [4.21]$$

where Y_{EC} is the ear canal admittance, which must be calculated from the models of the cochlea, middle ear and ear canal. The ear canal admittance can be related to the stapes impedance, Y_{St}, by manipulating [4.20] to give:

$$Y_{EC} = \frac{T_{22} Y_{St} - T_{12}}{T_{11} - T_{21} Y_{St}} \qquad [4.22]$$

Thus, in addition to the cochlear model, the model becomes fully specified (at any given stimulus frequency) when the six additional complex numbers defining the front end model are specified. These are: Q_{Src}, Y_{Src}, and the four terms in T_{ij}.

With this full model, the boundary condition given in [4.5] is no longer directly applicable, since the stapes velocity is not explicitly defined. Instead, specifying the volume velocity at the source, Q_{Src}, leads to a system of simultaneous equations which must be incorporated into the matrix equation representing the finite difference approximation to [4.4]. The system comprises equations [4.21] and [4.22], plus the following relationship for the stapes admittance in terms of the semi-difference pressure and semi difference pressure gradient, both evaluated at the base of the cochlea:

$$Y_{St} \equiv Q_{St} \,/\, p_{St}$$

where

$$Q_{St} \equiv A_{St} u_{St} \; ; \qquad u_{St} = -\frac{1}{i2\omega\rho} \frac{dp_d}{dx}\bigg|_{x=0} \qquad [4.23]$$

$$p_{St} = 2 p_d \,\big|_{x=0}$$

and where A_{St} is the area of the stapes.

Three front end models have been used in this thesis. The first is a perfectly transparent middle ear coupled to an OAE probe with a purely resistive admittance, which leads to negligible multiple reflection of the cochlear TW (discussed in section 4.9). This is similar to that used by Kanis and de Boer (1993b). The model is specified by $\boldsymbol{T}_{11} = 56$, $\boldsymbol{T}_{22} = 1/56$, $\boldsymbol{T}_{12} = \boldsymbol{T}_{21} = 0$, and $\boldsymbol{Y}_{\mathrm{Src}} = 2.941\ \mathrm{m^4skg^{-1}}$ for all stimulus frequencies. The value of $\boldsymbol{Y}_{\mathrm{Src}}$ was chosen to match the typical admittance seen looking into the ear canal, between 1 and 2 kHz. This ensures that, when a backward TW in the cochlea encounters the stapes it is largely dissipated by the source admittance, rather than being reflected. This model will be referred to as the 'ideal front end model'.

The second front end model has an identical transparent middle ear to the ideal front end model, but is coupled to an OAE probe with zero admittance, leading to perfect basal reflections (discussed in section 4.10). The model is specified by $\boldsymbol{T}_{11} = 56$, $\boldsymbol{T}_{22} = 1/56$, $\boldsymbol{T}_{12} = \boldsymbol{T}_{21} = 0$, and $\boldsymbol{Y}_{\mathrm{Src}} = 0$ for all stimulus frequencies. This model will be referred to as the 'high reflection front end model'.

The third front end model is a more complex and realistic model due to Kringlebotn (1988), in which the values of $\boldsymbol{T}_{ij}$ are strongly frequency dependent. The OAE probe is treated as a perfect volume velocity source, achieved by setting the admittance, $\boldsymbol{Y}_{\mathrm{Src}} = 0$. This gives a strongly reflective termination of the cochlea at the stapes. Note, however, that the stapes is still not a perfect reflector of backward cochlear TWs, because of dissipative elements in the middle ear which are included in the values of $\boldsymbol{T}_{ij}$. This model will be referred to a the 'Kringlebotn front end model'.

4.6.6 Numerical Solutions to the Longwave Model

The method of solution of equation [4.4] chosen in this thesis is the finite difference method outlined by Neely and Kim (1986). Because of its apical and basal boundary conditions, it is a boundary value problem, which must therefore be solved via simultaneous equations. Typically the CP is discretized into about 1000 - 2000 points, leading to a matrix equation that can be solved in seconds on a Pentium II PC to give the ear canal impedance for any given single frequency.

The quasilinear models are solved using an identical method to the linear models. They are, however, much slower because many iterations are required (at any given pair of primary stimuli) before the response converges. In addition, at each point on the CP, and for each iteration, it is necessary to convert the response into and out of the time domain, thus increasing the computational burden substantially.

4.7 Some Important Features of Cochlear Models and their Responses

4.7.1 The Wave Nature of the Cochlear Response

Equation [4.4] would reduce to the simple wave equation if the CP impedance were purely stiffness controlled (i.e., the stimulus frequency were well below resonance), and if the stiffness were varying only very slowly with place. This situation is approached near the base of the cochlea for low frequency stimuli. This resemblance to the wave equation also illustrates the wave bearing nature of the cochlea. Mechanical wave propagation requires a spatial distribution of two types of energy store: kinetic and potential. In the cochlea, the kinetic energy is stored in the motion of the cochlear fluid, whilst the potential energy is stored in the stiffness of the CP. Note that at frequencies above the characteristic frequency at a given place, the impedance becomes mass-like rather than spring-like and thus no longer acts as a potential energy store. Consequently TWs no longer propagate in this region[23] (Patuzzi, 1996).

There are a number of important differences between the equation [4.4] and the simple 1-D wave equation familiar from acoustics, in which the phase speed is both real and independent of frequency and place. In equation [4.4] the phase speed, c_{TW} is a function of both x and ω, and is a complex number which may lie in any one of the four quadrants of the complex plane. This gives rise to very complicated TW behaviour, even for this simplified 1-D, linear model. In fact, in the general case of equation [4.4], where the phase speed term, c_{TW}, is allowed to be any arbitrary function of x, the mechanical response of the cochlea, as defined by the solution, $p_d(x)$, of [4.4], cannot always be represented simply as the sum of forward and backward going "waves". In these cases, the definition of terms such as "wave" and "reflection" can become complicated. The general problem is discussed by Kaernbach *et al.* (1987), Shera and Zweig (1991a, 1991b, 1993a), de Boer and MacKay (1980), de Boer *et al.* (1986), Viergever and de Boer (1987), and Talmadge *et al.* (1998).

One important result arising from these analyses is as follows. Cochlear models which have no scattering impedance are approximately "reflectionless" in the sense that no ripples appear in the frequency response function of the driving point acoustic impedance

[23] The TW is said to be "evanescent" in this region.

measured at the oval window[24]. This result holds whether or not the cochlear amplifier is present, provided that the CP impedance varies only slowly with place (as is the case with the model parameters specified in section 4.6). This allows SFOAEs from the cochlear models to be defined in terms of the difference in the response measured in the ear canal between a reflecting and a reflectionless cochlear model. At first sight this may appear to differ from the definition of OAEs in terms of their origin in active cochlear processes (section 2.1). However, as is discussed in sections 3.4 and 5.3, it turns out that the cochlea is approximately reflectionless if either the scattering impedance or the cochlear amplifier impedance is zero.

A second useful result that arises when the scattering impedance is zero and when the spatial variation in the CP impedance[25] is slow is that a simplified method of solution of the wave equation [4.4] can be used. This method, known as the Wentzel-Kramers-Brillouin (WKB) method, obviates the need for the finite difference solution method described earlier (Zweig *et al.*, 1976; Shera and Zweig, 1991b; de Boer, 1996; Talmadge *et al.*, 1998; Viergever and Diependaal, 1986). It has been shown that the WKB method is applicable to most active and passive cochlear models over a wide range of stimulus frequencies (provided the scattering impedance is zero) (Shera and Zweig, 1991b). The fact that the WKB method is valid also provides insight into the nature of the mechanical response since it allows the response to be written as the sum of forward and backward going waves. This further illustrates the wave bearing nature of the cochlea and allows the flow of wave energy within the cochlea to be easily identified. The validity of the WKB method implies that the cochlea is reflectionless[26] as discussed above. However, the reverse is not necessarily true: the cochlea may be reflectionless without the WKB method being applicable (Shera and Zweig, 1991b). The WKB method has been used in this thesis to check the results of the finite difference method, but has not been used for the final results since it is not directly applicable when the scattering impedance is non-zero.

It is sometimes mistakenly reported (e.g., Pickles, 1988, p. 49) that backward TWs are prohibited by the spatial variation in the BM properties. This appeared to be demonstrated by an experiment in which acoustic stimulation at the apex of a cat cochlea appeared to generate a forward TW originating at the stapes, rather than the expected backward TW (Lighthill, 1981, p. 178). A full explanation for this apparently paradoxical

[24] Contradicting this general result, Talmadge *et al.* (2000) present results from a 'smooth' cochlear model which show some reflections arising purely from cochlear nonlinearity. However, these nonlinear reflections give rise to ripple with very high frequency spacing (i.e., high periodicity), and to ripple amplitudes that are generally much less than those due to the scattering impedance.
[25] More strictly the wave impedance, which is related to the CP impedance, and the height of the scalae.
[26] When the WKB method applies, reflections can only arise at the boundaries of the CP.

result is given by Lighthill (1981). Qualitatively, this argument is most easily understood by considering a volume velocity excitation of the apical wall of the cochlea which is symmetric in the upper and lower channels[27]. Symmetric excitation (i.e., push-push) does not couple to the cochlear TW, but instead generates a 'fast' compression wave that travels backward (in both channels) to the base of the cochlea. Here it encounters asymmetric boundary conditions due to the different impedances of the oval and round windows. This gives rise to asymmetric velocities in the upper and lower channel, which in turn give rise to a forward cochlear TW, as observed (recall that only asymmetric loading couples to the TW). Even if the apical excitation is not purely symmetric, any symmetric component will still generate a significant fast wave followed by the forward TW. Also, any asymmetric component will only initiate a significant backward TW if the stimulus frequency is lower than the characteristic frequency at the cochlea apex. Therefore it is easier to generate forward cochlear TWs than backward ones using an apical volume velocity source. Note that this situation does not arise in the case of the cochlear amplifier. To see this, consider the hypothetical case in which the cochlear amplifier applies a force at a single point on the CP, in the absence of any other excitation. Unlike the previous case, the alternating force on the CP leads automatically to a purely asymmetric loading of the two channels. This then generates both forward and backward TWs radiating out from the point of application of the force. Because there is no symmetric loading, there is no fast wave.

4.7.2 Validity of the Longwave Model

The longwave approximation becomes applicable when the wavelength is much longer than the height of the channel (de Boer, 1996, suggests $kH<1$). In this case the fluid y-momentum equation may be ignored (as it is in the derivation of the longwave equation).

The longwave approximations break down as the wavelength of the TW becomes much smaller than the height of the cochlear channel, which occurs in the region of the peak of the TW envelope. Since it is in this region that SFOAEs are thought to be generated, this may seem like a fatal shortcoming of the longwave model for predicting SFOAEs. However, it is claimed by Zweig and Shera (1995) that the longwave model qualitatively captures all the features of the TW that are essential for SFOAE generation. The main feature is the 'tall and broad' TW envelope. A more general argument for the use of the longwave approximation in macromechanical models is given by Zweig (1991,

[27] Equally a volume velocity source could be introduced.

p. 1246). A discussion of the use of the longwave model for the analysis of self-suppression, two-tone suppression and DPOAE generation in macromechanical models is given by Kanis and de Boer (1993b, 1994, 1997).

4.7.3 Higher Dimensional Models

Several 2 and 3-dimensional cochlear models have been developed, which have the advantage over 1-D models of more accurately modelling the shape of the TW peak. These are discussed by Lighthill (1981), and de Boer (1980, 1996). These could perhaps be useful for improving the quantitative predictions of SFOAEs. However, following Zweig and Shera (1995), and avoiding excessive complexity, this possibility has not been investigated in this thesis.

4.7.4 Active Processes and Stability

In active cochlear models, the cochlear amplifier is included in equation [4.4] as a (linear) component in the CP point impedance, and is of crucial importance to the generation of SFOAEs. Two questions are sometimes asked. Firstly, does the cochlear amplifier provide 'undamping' or 'amplification'? Secondly, is the cochlea stable (de Boer; 1993)?

'Undamping' means that the active processes act to oppose the passive damping, but that overall the CP remains positively damped (i.e., the real part of Z_{CA} is negative in some regions of the place-frequency plane, but the real part of Z_{CP} always remains positive). The predicted response of an active system with undamping is no different from a passive system with light damping. It has been argued that the measured cochlear mechanical response cannot be explained by undamping alone, but that active amplification is required (de Boer, 1983, 1993; Zweig, 1991). Also, SOAEs clearly cannot arise from a system with undamping.

A system with active amplification is one which appears to be negatively damped (i.e., the real part of Z_{CP} becomes negative) in certain regions of the place-frequency plane. Such a system is capable of injecting energy into the TW to enhance it. As discussed previously, this does not, by itself, generate OAEs. Instead it creates a region of the CP in which the TW grows, but is not reflected. This is the reverse of the situation where a wave propagates through a dissipative system: the amplitude of the wave decays, but no reflection occur.

In systems with active amplification, the problem of instability may arise, which might have undesirable effects on the performance of the cochlea. Two aspects of instability have been distinguished in the literature. Firstly a point on the CP, when isolated from the rest of the model, may be unstable. This can simply be checked by examining $\mathbf{Z}_{CP}$ as a function of frequency at each point and applying the usual stability criteria for a transfer function (Hsu, 1995). Secondly, and more importantly, the entire system may be unstable. This can only be checked by solving equation [4.4], subject to its boundary conditions. Because these depend on the middle ear model and on the impedance of the OAE probe, so too will the stability of the cochlea. It is this second form of instability that is the more important one for this thesis.

The CP impedance of each element used by Kanis and de Boer (1993b) was unstable when in isolation, but became stable when all the elements were coupled together in their model. This is because, for any given frequency, there is negative damping over only a limited region of the CP. Elements within this region become stabilised by their hydromechanical coupling to positively damped regions. Also, Kanis and de Boer (1993b) deliberately chose a middle ear model that would prevent instabilities, by minimising any basal reflection of the TW.

The CP impedance suggested by Zweig (1990, 1991) is stable both when in isolation (despite having a negative resistance at certain place-frequencies), and when coupled together within the cochlea (except at certain frequencies, as discussed below).

Use of a simple negative damper (i.e., overall negative resistance at all places and all frequencies) would lead to both instability of the elements in isolation, and when coupled together in the cochlea. This type of activity is included in the model reported by van Hengel *et al.* (1996) where strong nonlinearity is proposed to limit the instability. The problem of instability in models is discussed by Koshigoe and Tubis (1983) and Zweig (1991).

Cochlear models including active processes of the form given by Kanis and de Boer (1993b) or Zweig (1991), though stable at most frequencies, can become unstable at certain frequencies when inhomogeneities are introduced leading to reflection of the forward TW. These instabilities, arising from multiple reflections due to the CP inhomogeneities and the impedance mismatch at the stapes, are thought to be the cause of SOAEs (e.g., Kemp, 1980; Talmadge and Tubis, 1993; Talmadge *et al.*, 1998). The precise values of these frequencies depend on the impedance seen by the TW "looking out" of the oval window, and are therefore dependent on the middle ear, ear canal and OAE probe impedances.

4.7.5 Scaling Symmetry

Scaling symmetry arises when all the variables in equation [4.4] are functions not of x and ω independently, but of a single variable, $\beta_{x\text{-}\omega}$, defined here as:

$$\beta_{x-\omega}(x,\omega) = \frac{\omega}{\omega_{c_0} e^{-x/d}} \qquad [4.24]$$

(which defines the ratio of the stimulus frequency, ω, to the local natural frequency, $\omega_{c_0} e^{-x/d}$, where ω_{c_0} is the basal natural frequency). This means that cochlear responses (such as the fluid pressure or CP velocity) can be represented as a one dimensional curve against the β-axis instead of as a 2-D surface above the x-ω plane. For this to occur, the CP impedance must vary with x and ω in a particular way (Shera and Zweig, 1991a). Note that the passive impedance in equations [4.9] and [4.10] is an example of one leading to approximate scaling symmetry. (Even greater scaling symmetry arises when the cross-section of the cochlear scalae tapers exponentially; Shera and Zweig, 1991a). Scaling symmetry in the cochlea has two important consequences. The first is that the cochlear response shows symmetries that greatly simplify the analysis. The second is that no SFOAEs with spectral periodicity are generated.

The first consequence is that the spatial pattern of the pure-tone cochlear response shows shift symmetry. This means that the pure-tone TW magnitude and phase responses (i.e., spatial pattern) at two different stimulus frequencies are identical, except for a translation in place. It follows then that knowledge of the pure-tone TW response along the cochlea at a single stimulus frequency is sufficient to define the response along the cochlea at all frequencies. Alternatively, knowledge of the frequency response at any fixed point on the cochlea is sufficient to define the frequency response at all places. Furthermore, place and frequency show a very simple symmetry. This symmetry means, for example, that the graph of the CP velocity plotted against cochlear place at a fixed stimulus frequency looks identical to the graph of CP velocity plotted against the logarithm of stimulus frequency at fixed place. Thus cochlear place and the logarithm of stimulus frequency become completely interchangeable (requiring only multiplication by a simple scaling factor).

The second consequence is that the phase response of any SFOAE from a scaling symmetric cochlea is virtually flat with frequency. This is unlike the steep phase curve which is approximately linear with logarithmic frequency, and which is observed experimentally. This result led Kemp (1986) to conclude that the observed SFOAEs could

not arise from a 'wave-fixed' reflection site. This result is further discussed by Strube (1989), Zweig and Shera (1995) and Shera and Guinan (1999).

In the theory of Zweig and Shera (1995), the cochlea shows only partial scaling symmetry. They argue that, to a first approximation, the cochlea shows scaling symmetry with regard to its primary response (i.e., the shape of the forward TW). However, place fixed features in the cochlea cause a break from scaling symmetry leading to lead to secondary responses such as SFOAEs. These secondary responses are, in general, significantly less than the primary responses. To achieve this situation in their cochlear models, the passive and active CP impedances in equation [4.8] show approximate scaling symmetry, whilst the scattering impedance does not. Thus, the scattering impedance provides the break from scaling symmetry which leads to the generation of realistic (long latency) SFOAEs in the first place, whilst the passive and active impedances show the approximate scaling symmetry which leads the scattering sites to generate SFOAEs of approximately constant periodicity.

4.7.6 Alternative Formulations of Published Cochlear Models

In this section, the cochlear models used in this thesis are compared with other models found in the literature. Often the longwave equation [4.4] is implemented as a discrete 1-D transmission line, where the cochlear fluid mass and the CP mechanics are replaced by their electrical analogues. These transmission line models are essentially the same as the discrete form of the wave equation described above. A comparison of several transmission line models demonstrating their high degree of equivalence is given by de Boer (1995).

Models have been developed with several different objectives in mind. One objective has been to match those cochlear responses that have been measured directly in experiments, such as BM vibration, or hair cell potentials, or neural responses (e.g., Zwicker, 1979; Neely and Kim, 1986; Geisler, 1991; Kolston, 1988; Kolston and Smoorenburg, 1990; Neely, 1993; Zweig, 1990, 1991; Allen, 1988, 1990; Kanis and de Boer, 1993a). Some of these models include a cochlear amplifier in some form, and also a more detailed micromechanical model, but they all exclude the scattering impedance required to predict SFOAEs.

Kanis and de Boer (1997) included a nonlinear cochlear amplifier which leads to the generation of DPOAEs from the model, but exclude a scattering impedance, and therefore are unable to generate SOAEs, SFOAEs or TEOAEs.

Of more direct relevance are models that include both a cochlear amplifier and a scattering impedance. These include models reported by Furst and Lapid (1988), Fukazawa (1992), Fukazawa and Tanaka (1996), Zwicker and Lumer (1985), Zwicker and Peisl (1990), Zweig and Shera (1995), Talmadge *et al.* (1998) and Wada *et al.* (1999). The main results of these models can be reproduced using the model developed here. These are presented in section 5.3.

4.8 Zweig and Shera's Phenomenological Model of SFOAEs

Zweig and Shera (1995) present a simplified analysis of the 1-D longwave model, which gives a useful approximate relationship between the reflectance measured at the base of the cochlea, the shape of the TW and the CP scattering impedance. This relationship holds for a large class of scattering impedance functions[28], not just the random function proposed by Shera and Zweig. In this thesis, this simplified model has not been used to generate any of the final quantitative predictions in this thesis. However, it is introduced here because of its power in explaining the basic form of SFOAEs.

The basal reflectance (introduced in section 2.2) relates the backward TW to the forward TW at the base of the cochlea. Being evaluated at the base of the cochlea, it contains information about both the scattering mechanism and the TW propagation to and from all the scattering sites apical to the base. The phase of the reflectance is related to the overall delay arising from TW propagation as well as to any phase changes induced by scattering. The magnitude of the reflectance is related to the amplification or dissipation of the TW occurring during propagation, as well as to the degree of backward scattering. As seen in fig. 2.4 and 2.5, the form of this basal reflectance closely resembles the SFOAE measured in the ear canal. In fact the two are related by a fairly simple equation given in the next section.

In the phenomenological model, it is first assumed that the primary cochlear responses exhibit scaling symmetry, which means that it is possible to transform cochlear location directly onto a stimulus frequency scale (as discussed in section 4.7.4). Second, it is assumed that the scattering impedance, $\boldsymbol{Z}_{Sc}$, is roughly independent of frequency (as it would be if the inhomogeneities were only in the CP damping). Zweig and Shera then show that the relationship between the cochlear reflectance, $\boldsymbol{R}$, the CP scattering

[28] For this relationship to hold, the scattering impedance must be independent of stimulus frequency, but may have any arbitrary spatial variation.

impedance, $\boldsymbol{Z}_{Sc}$, and the TW function, $\boldsymbol{T}$, can be cast in the form of a simple linear filtering operation:

$$\boldsymbol{R}(\eta) = \rho_{Sc}(\eta) \otimes \boldsymbol{T}^2(\eta)$$

$$\eta(\omega) \equiv -\ln(\omega / \omega_{C_0}) \qquad [4.25]$$

$$\boldsymbol{T}\ (\Theta) \equiv \frac{v_{CP}(x,\omega)}{\boldsymbol{u}_{St}(\omega)}$$

$$\Theta(x,\omega) \equiv \ln\left[\frac{\omega}{\omega_{C0} e^{-x/d}}\right] = \frac{x}{d} - \eta(\omega)$$

where $\boldsymbol{R}$ is the cochlear reflectance; ρ_{Sc} is called the scattering potential, and is roughly proportional to $\boldsymbol{Z}_{Sc}$; $\boldsymbol{T}$ is the TW function defined as the basilar membrane velocity, v_{CP}, normalised with stapes velocity, $\boldsymbol{u}_{St}$; and $\otimes$ denotes convolution. Scaling symmetry ensures that $\boldsymbol{T}$ is a function of a single variable, Θ, rather than of both x and ω (Note that $\Theta = \ln \beta_{x\text{-}\omega}$ where $\beta_{x\text{-}\omega}$ is given in [4.24].) The independent variable in the filtering operation, η, is the natural logarithm of the stimulus frequency, normalised with a reference frequency as in [2.11]. This quantity is denoted by $\hat{\chi}$ in Zweig and Shera's notation (1995).

From equation [4.25], the mathematical recipe for generating the reflectance spectrum is as follows. To form $\rho_{Sc}(\eta)$, first plot the scattering potential as a function of cochlear place, x, then map place to the characteristic frequency using [4.10]. Plot this on a logarithmic frequency scale, denoted by η. To form $\boldsymbol{T}(\eta)$, plot the CP velocity response against place, x, at a single frequency, normalise with the stapes velocity, and again map place to the log frequency variable, η as before. To obtain $\boldsymbol{R}(\eta)$, $\rho_{Sc}(\eta)$ is passed through a filter whose impulse response function is given by $\boldsymbol{T}(\eta)$. Thus $\boldsymbol{R}(\eta)$ contains the spatial frequency components in $\rho_{Sc}(\eta)$, filtered by the spatial passband of the TW.

Physically the act of sweeping the stimulus frequency is to sweep the TW peak across the CP inhomogeneities. The forward TW in the peak region is reflected by spatial arrangements of the inhomogeneities that have a spatial period equal to half the TW wavelength. The forward TW outside the peak region is too small to give strong reflections.

Fourier transforming equation [4.25], gives:

$$F\{\boldsymbol{R}\} = F\{\rho_{Sc}\} F\{\boldsymbol{T}^2\} \qquad [4.26]$$

$$\eta \Leftrightarrow \phi$$

where $F\{\ldots\}$ denotes the Fourier transform, from the logarithmic frequency variable, η, to its conjugate, ϕ, as defined by:

$$F\{A\} \equiv \int_{-\infty}^{\infty} A(\eta)\exp(-i2\pi\phi\eta)d\eta \qquad [4.27]$$

For Shera and Zweig's model, $F\{\boldsymbol{R}\}$ appears pulse like in the ϕ-domain. The peak of the pulse occurs at a location called here the ϕ-centre value and denoted by ϕ_C. (Zweig and Shera (1995) denote this $\hat{\varphi}/2\pi$. The value of 2π arises from their definition of the Fourier transform, which maps time to radian frequency). The value of ϕ_C, is related directly to the 'average' single-ripple spectral periodicity, Ψ, of the SFOAE, by the equation (Zweig and Shera, 1995):

$$\Psi \equiv \frac{\Delta f}{f_{GMF}} \approx \frac{1}{\phi_C} \qquad [4.28]$$

where, in humans, typically $\Psi \approx 1/15$. In [4.28] Δf is the peak to peak frequency interval, and f_{GMF} is the geometric mean frequency of two adjacent peaks. It follows from [2.15] that ϕ_C is also equal to the group cycle delay, ϕ_{GCD}.

The following terminology will be adopted to distinguish the various different independent variables used in representing the SFOAE and related signals. The term 'frequency spectrum' indicates that the signal is tabulated as a function of linear frequency, f. An 'η-series' or 'η-function' indicates that the signal is tabulated as a function of the logarithmic frequency, η. The term 'ϕ-spectrum' indicates the signal has first been mapped to the η-domain, and then Fourier transformed to the ϕ-domain, by [4.27].

This phenomenological model allows the cochlear reflectance to be estimated for a given scattering potential and TW function. Fig 4.4 illustrates the form of the reflectance that results from the interaction between a random array of CP inhomogeneities, and a simplified TW function. This TW function represents the peak region only, and has been approximated by a simple pulse shaped envelope and a linear phase (i.e., a constant wavelength). The cochlear reflectance has been synthesised digitally in Matlab by generating a white-noise signal to represent the spatially random scattering potential. This is then passed through a band-pass filter representing the TW function. The resulting band-pass signal then gives the estimated cochlear reflectance as in [4.26]. In fig 4.4, cochlear location, x, and spatial frequency have been transformed into η and ϕ respectively.

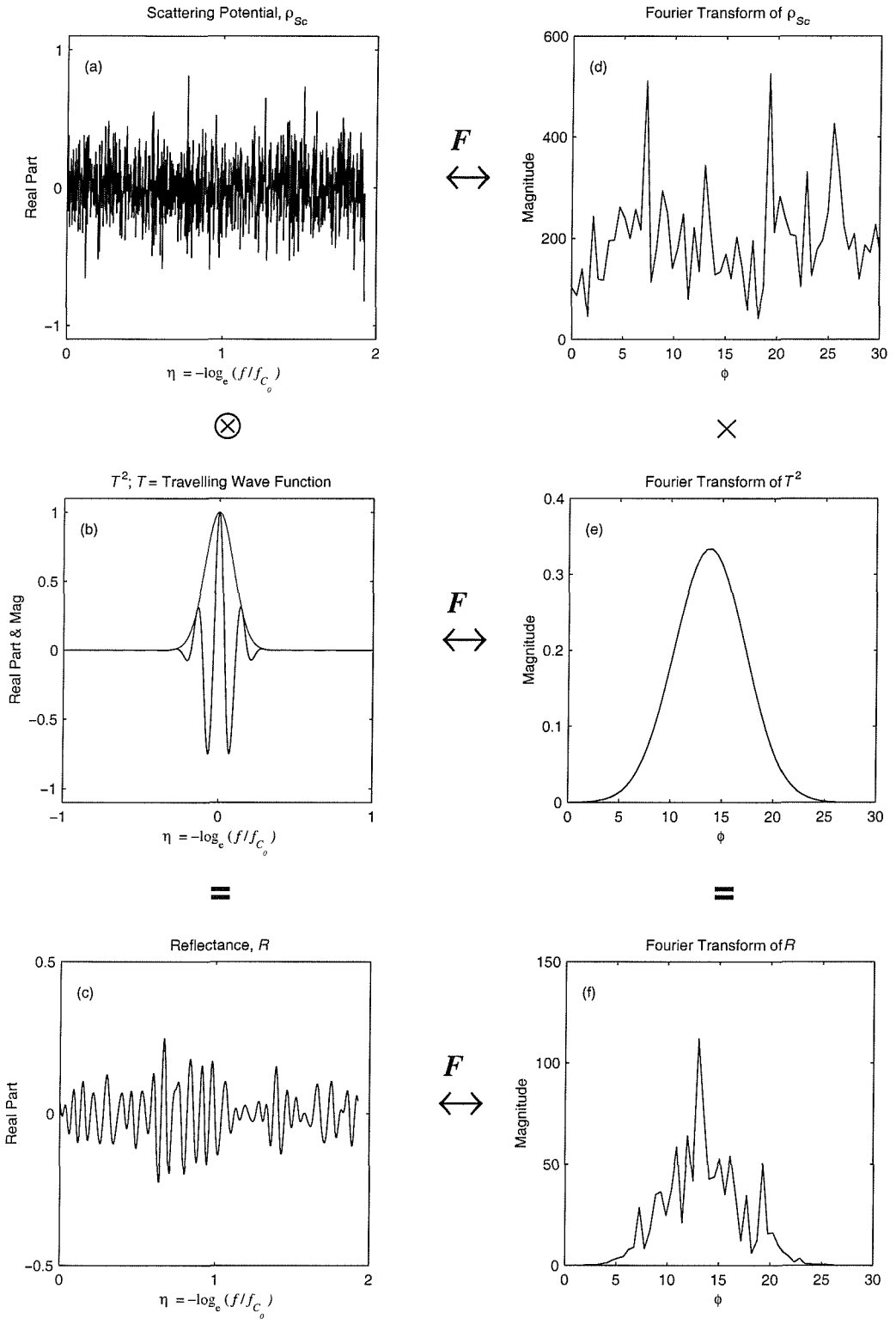


Fig 4.4. Illustration of the predicted cochlear reflectance due to spatially random inhomogeneities, according to Zweig and Shera's phenomenological model (1995). The cochlear reflectance has been synthesised from an assumed scattering potential and travelling wave function. The scattering potential is a broad band random function of place, with place transformed to η via the place-frequency mapping. A simplified travelling wave function with linear phase and gaussian envelope has been assumed. The reflectance is calculated from the convolution of the scattering potential with the travelling wave function in the η-domain, where η is the logarithm of the stimulus frequency, normalised with the characteristic frequency at the stapes. Panel (a): Scattering potential, modelled as a broad band random irregularity along the basilar membrane. Panel (b): simplified travelling wave function modelled as a single pulse of constant wavelength. Panel (c): resulting cochlear reflectance. Panels (d), (e) and (f) are the Fourier transforms of (a), (b) and (c) respectively, where ϕ is the conjugate Fourier variable.

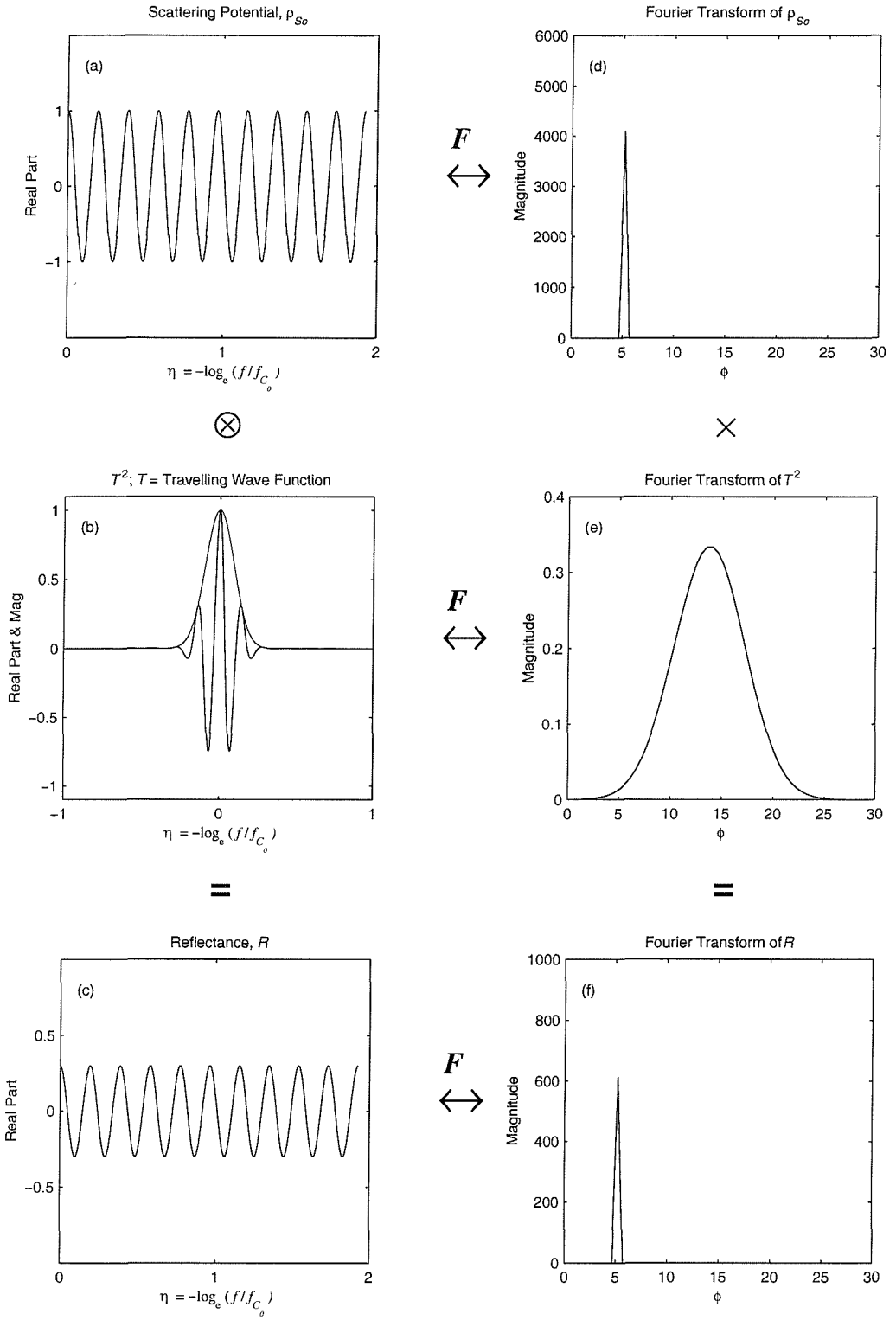


Fig 4.5. Illustration of the predicted cochlear reflectance due to spatially periodic inhomogeneities, according to Zweig and Shera's phenomenological model (1995). The cochlear reflectance has been synthesised from an assumed scattering potential and travelling wave function. The scattering potential is a periodic function of place, as suggested by Strube (1989), with place transformed to η via the place-frequency mapping. A simplified travelling wave function with linear phase and gaussian envelope has been assumed. The reflectance is calculated from the convolution of the scattering potential with the travelling wave function in the η-domain, where η is the logarithm of the stimulus frequency, normalised with the characteristic frequency at the stapes. Panel (a): Scattering potential, modelled as a broad band random irregularity along the basilar membrane. Panel (b): simplified travelling wave function modelled as a single pulse of constant wavelength. Panel (c): resulting cochlear reflectance. Panels (d), (e) and (f) are the Fourier transforms of (a), (b) and (c) respectively, where ϕ is the conjugate Fourier variable.

Fig 4.5 shows the synthesised reflectance resulting from a Strube type CP corrugation interacting with the same simplified TW function. This was synthesised in Matlab as before except that a sine wave is used in place of the white noise signal to represent the CP scattering potential.

These two figures illustrate several important points. The ϕ-domain is related to the spatial domain by the Fourier transform. Thus central location of a function in the ϕ-domain (which will be referred to as the ϕ-centre value of the function) corresponds to a spatial frequency, whilst the width of the function in the ϕ-domain (which will be referred to as the ϕ-bandwidth of the function) is inversely related to the width of the function in the spatial domain. In the Shera and Zweig model, the periodicity in the reflectance (and therefore in the SFOAE frequency function) is determined predominantly by the location in the ϕ-domain of the peak in the Fourier transform of the TW function. This peak location is determined by the wavelength. (Notice, however, that the peak of the reflectance in the ϕ-domain does not coincide exactly with the peak in the TW in the ϕ-domain, because of the random nature of the scattering potential.) In contrast, in the Strube model, the periodicity is determined by the location in the ϕ-domain of the peak in the Fourier transform of the scattering potential function, which is determined by the spatial period of the cochlear inhomogeneities. Thus in Shera and Zweig's model it is the TW wavelength which determines the periodicity, whilst in Strube's it is the corrugation wavelength. This is important because it means that according to Shera and Zweig's model, but not Strube's model, changes in the TW function will show up in the SFOAE frequency function. For example, increasing the level of the input signal is known to cause the (normalised) TW envelope to reduce in height, to broaden, and for the wavelength increase. For Shera and Zweig's model, these three changes will manifest themselves in the ϕ-domain of the SFOAE as a reduction in height, a narrowing of width[29] and a shift in the peak towards smaller values of ϕ. This thesis sets out to look for experimental confirmation of these effects predicted by Shera and Zweig.

It could be argued that the SFOAEs arising from Strube's model (fig. 4.5) are unrealistically regular. However, this is not a fundamental objection to Strube's theory. More realistic SFOAEs would have been obtained in fig. 4.5 if the spatially periodic scattering potential (fig. 4.5a) were replaced with a narrow-band random scattering potential. This model could then be thought of as lying somewhere between Strube's proposed periodic scattering potential and Shera and Zweig's broad band random scattering

[29] This follows from the 'inverse spreading' relationship which holds between a pulse-shaped function and its (pulse-shaped) Fourier transform (e.g., Randall, 1997).

potential. However, there remains an essential difference between the two theories: in Strube's theory the periodicity is determined predominantly by the spatial periodicity in the CP scattering potential (i.e., the input to the spatial filter), whereas in Shera and Zweig's model the periodicity arises from the TW wavelength near the peak of the TW function (i.e., the centre frequency to the spatial filter). From consideration of Shera and Zweig's phenomenological model, it can be seen that, provided the ϕ-bandwidth of the CP scattering impedance function is significantly less than that of the TW function, and that its ϕ-centre value falls within the pass-band of the TW 'spatial filter' then the resulting SFOAE periodicity will be dominated by the spatial periodicity of the CP scattering potential, as suggested by Strube. Thus a model may be described as a Strube model whenever the SFOAE periodicity is dominated by the spatial periodicity of the CP scattering potential, rather than by the wavelength at the TW peak.

The form of the cochlear reflectance and of the SFOAE predicted by Shera and Zweig's phenomenological theory with spatially random inhomogeneities is further illustrated in fig. 4.6. For ease of interpretation in this (and subsequent) illustrations of signals in the η-domain, the value of η will be converted by a linear transformation into the more familiar units of octaves, denoted η_{Oct} as in [2.10]. These are calculated with respect to an arbitrary reference frequency, f_{Ref}, chosen here to be 1 kHz.

$$\eta_{Oct} \equiv \log_2(f / f_{Ref})$$

$$\therefore \eta_{Oct} = \frac{-\eta}{\ln(2)} + A \qquad [4.29]$$

Note that a factor of about 0.69 is introduced in converting η to η_{Oct}. Thus, if the function when plotted against η_{Oct} exhibits about 10 ripple cycles per octave, (i.e., a ripple period of 1/10 octaves) then the ϕ-spectrum will show a peak at around $10/0.69 \approx 15$, indicating a periodicity of 1/15. Note that the conversion to octaves has not been made in figs. 4.4 and 4.5, since these plots emphasise the fundamental filtering relationships, where the change of variable would be confusing.

$$A = \frac{\ln(\omega_{C_0} / 2\pi f_{Ref})}{\ln(2)} \cdot \ln(2) = 0.6931$$

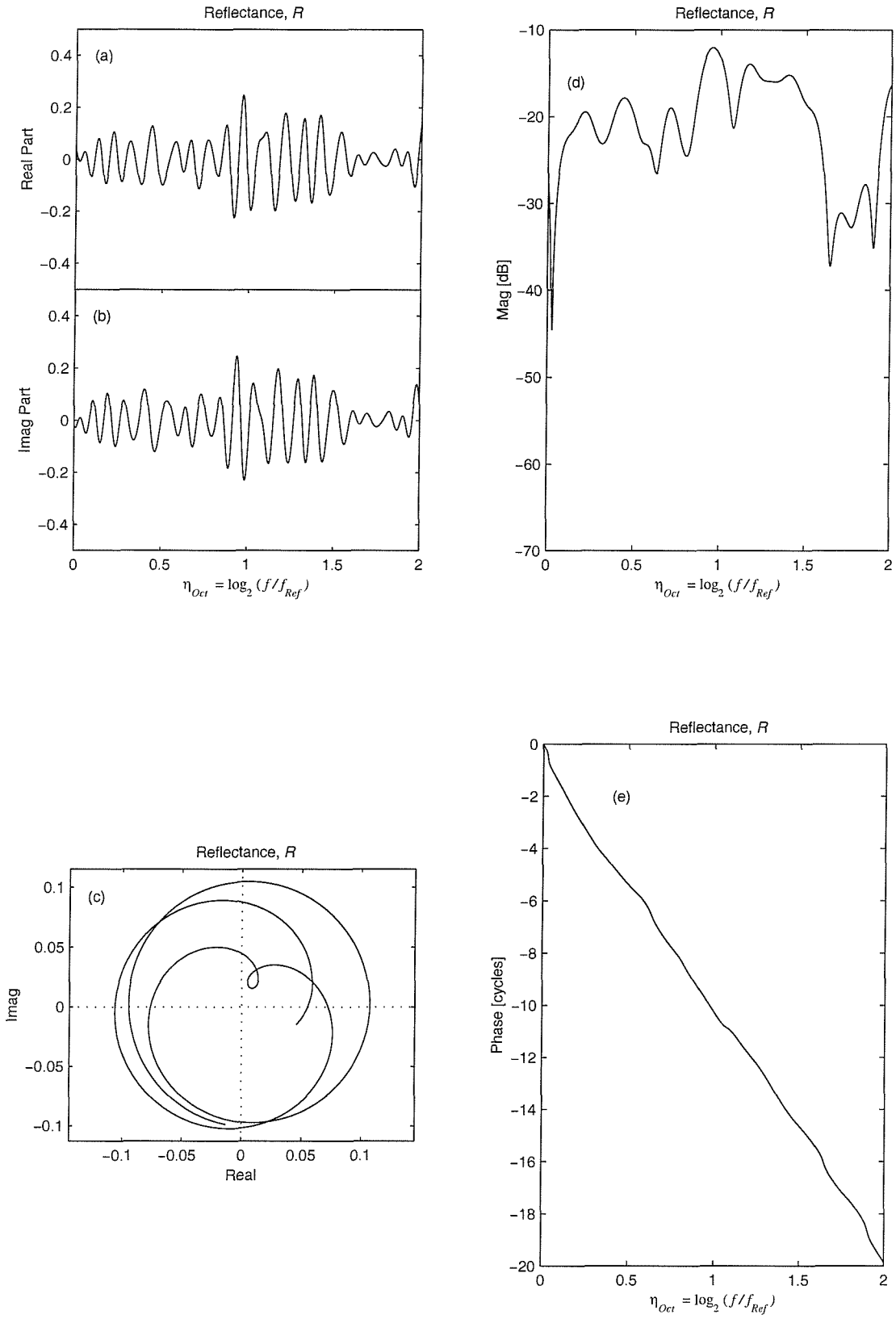


Fig 4.6 Representations of the complex cochlear reflectance synthesised using Zweig and Shera's phenomenological model (1995) with a random scattering potential. Panels (a) and (b): the real and imaginary parts plotted against the η_{Oct} (the octave form of the logarithm of the normalised stimulus frequency). Panel (c): real part plotted against the imaginary part. Panels (d) and (e): the magnitude and phase of the reflectance plotted against the η_{Oct}-variable.

Fig 4.6 shows how the periodicity appears in several alternative representations of the cochlear reflectance, $\boldsymbol{R}$, in the η (or log frequency) domain, as predicted by Shera and Zweig's phenomenological theory. According to this theory, $F\{\boldsymbol{T}^2\}$ is one sided (i.e., it is non-zero only for positive values of ϕ). Consequently, from equation [4.26], $F\{\boldsymbol{R}\}$ is also one sided, and therefore the function $\boldsymbol{R}(\eta)$ has a special form known as the analytic form, whereby the real and imaginary parts of $\boldsymbol{R}$ form a Hilbert transform pair. As has already been discussed in section 2.13, this means that in the η-domain, the real part of $\boldsymbol{R}$ resembles the imaginary part delayed by 90°, as is seen in figs 4.6 (a) and (b). The periodicity of the SFOAE, which is here approximately 1/15, shows itself as a ripple in the real and imaginary parts in figs 4.6 (a) and (b) with a peak-to-peak interval, $\Delta\eta_{Oct}$, of about 1/10. Fig 4.6 (c) shows how, as the stimulus frequency increases, the locus of $\boldsymbol{R}$ loops around the origin of the complex plane, at a rate of one cycle per increment in η_{Oct} of 1/10. Figs 4.6 (e) and (f) show that the magnitude of the reflectance is typically around 0.2, while the slope of the phase is roughly constant, at about one cycle per increment in η_{Oct} of 1/10. The periodicity does not show up at all in the magnitude of $\boldsymbol{R}$ shown in fig. 4.6 (d). The random fluctuations seen here are determined by the ϕ-bandwidth rather than the ϕ-centre value. The cochlear reflectance in fig. 4.6 resembles the idealised OAE (in the absence of multiple reflections) shown in fig. 2.4, except that its magnitude fluctuates randomly with frequency.

4.9 Relating the Ear Canal Pressure and the SFOAE Frequency Spectrum to the Cochlear Reflectance

The relationship between the SFOAE pressure, p_{SF}, at the probe microphone and the cochlear reflectance, $\boldsymbol{R}$, has been derived by Kemp (1980) and Shera and Zweig (1993a). This analysis uses a two-port network model of the middle and outer ear, coupled to the cochlear input impedance which is characterised in terms of the cochlear reflectance and non-reflecting impedance. The result from Shera and Zweig (1993a) is given in equation [4.30]:

$$p_{\Delta} = \frac{g\boldsymbol{R}}{1-r\boldsymbol{R}} \approx g\boldsymbol{R}(1+r\boldsymbol{R}+(r\boldsymbol{R})^2+\ldots)$$

$$p_{\Delta} \equiv \frac{p_{SF}}{p_{EC:R=0}} \qquad [4.30]$$

$$p_{SF} \equiv p_{EC} - p_{EC:R=0}$$

Here p_{EC} is the ear canal pressure measured for a cochlear reflectance of R whilst $p_{EC:R=0}$ is the ear canal pressure that would have been measured if the cochlear reflectance were zero. The SFOAE pressure, p_{SF}, is defined as the difference between these two (complex) pressures. The non-dimensional quantity, p_{Δ}, which will be referred to as the 'normalised SFOAE', is obtained by dividing the SFOAE pressure, p_{SF}, with the zero-cochlear-reflectance ear canal pressure, $p_{EC:R=0}$. The functions g and r are related to the middle ear, ear canal, and probe impedance characteristics. They can be calculated from the front end model parameters given in section 4.6.5. All terms are complex functions of frequency. The expansion for p_{Δ} shows how the SFOAE can be represented as the sum of multiple reflections within the cochlea (cf. equation [2.4]). The quantity, r (termed the basal reflectance) is the reflectance at the stapes as seen by a backward TW leaving the cochlea, whilst R (the apical cochlear reflectance) is the reflectance due to the CP scattering inhomogeneities encountered by a forward TW. Therefore the forward wave of amplitude A returns to the stapes as a backward wave of amplitude AR, which is further partially reflected at the stapes to give a new forward wave of amplitude ArR and so on. Each time the backward wave encounters the stapes, a proportional of the energy is transmitted out into the ear canal. These multiple backward waves sum in the ear canal to give the measured OAE. As for the idealised OAE considered in fig. 2.5, the magnitude of R determines the size of the first reflection, whilst the product rR determines how significant multiple reflections are.

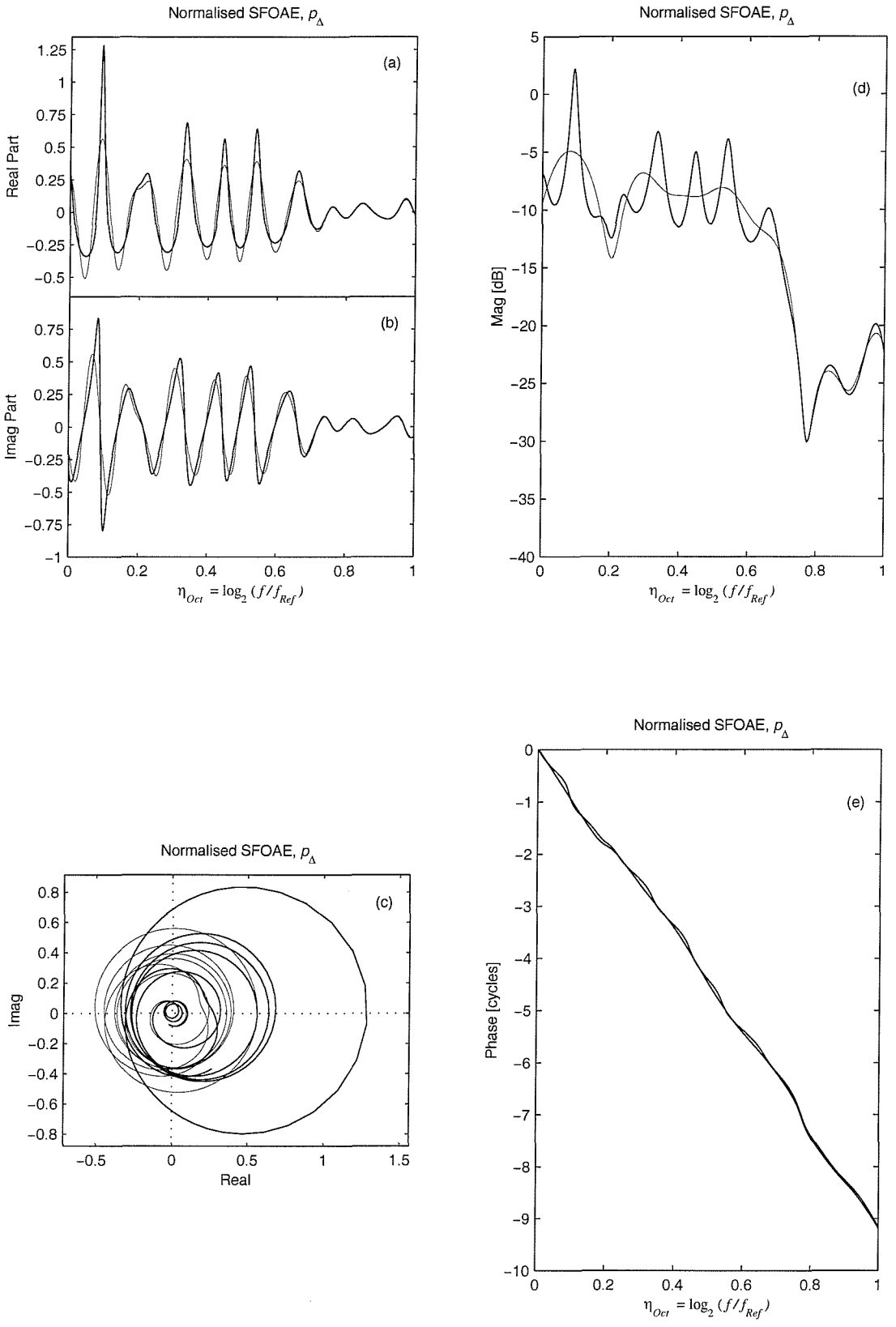


Fig 4.7. Representations of the complex normalised SFOAE pressure synthesised using Zweig and Shera's phenomenological model (1995) with a random scattering potential, and assuming that $\boldsymbol{g} = 1$; $\boldsymbol{r} = 1$ in equation [4.30]. Panels (a) and (b): the real and imaginary parts plotted against the η_{Oct} variable. Panel (c): real part plotted against the imaginary part. Panels (d) and (e): the magnitude and phase of the reflectance plotted against the η_{Oct}-variable. Thick line = complex normalised SFOAE pressure; thin line = cochlear reflectance.

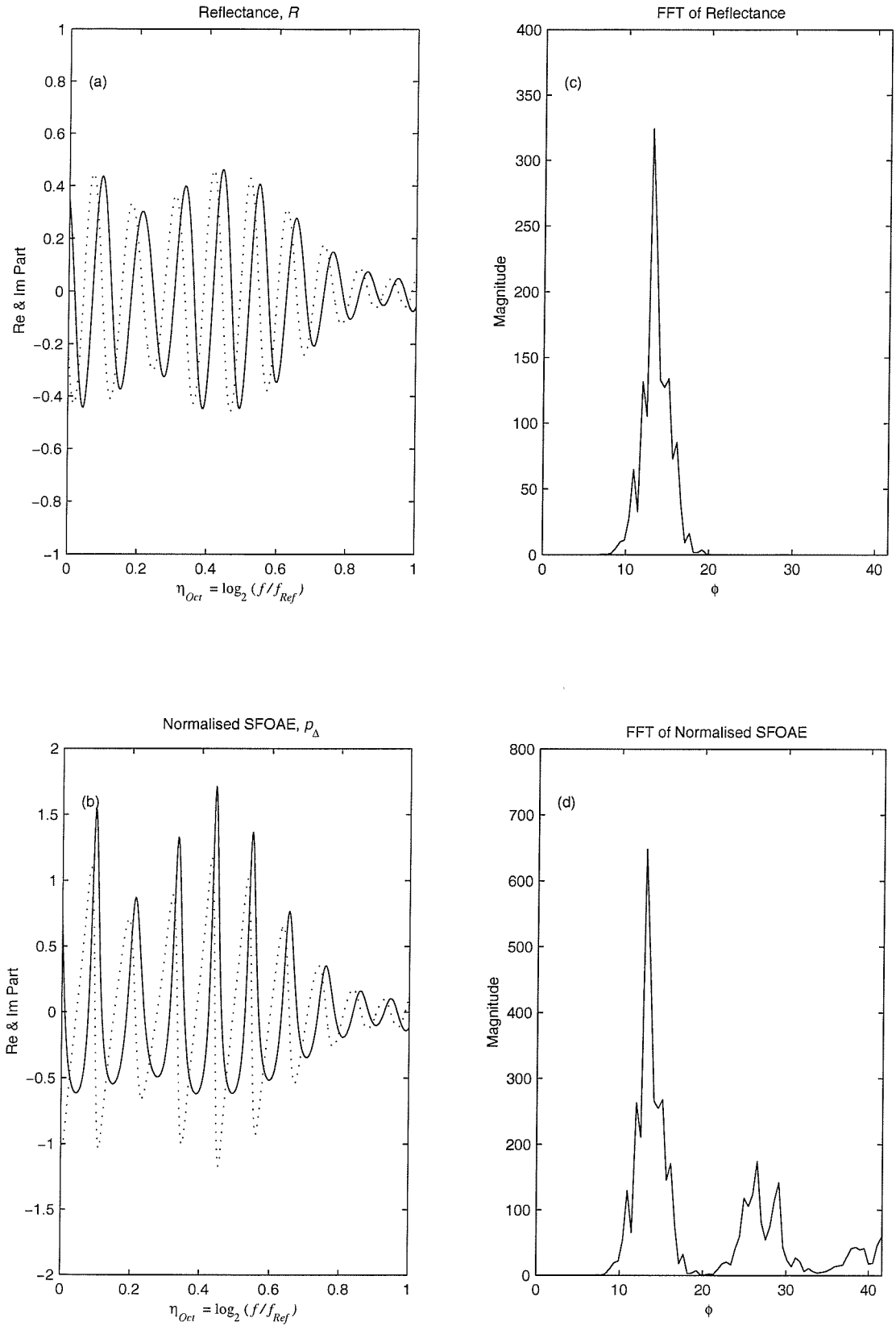


Fig 4.8 The effect of multiple reflections due to the middle ear on the predicted normalised SFOAE pressure in the η_{Oct} and ϕ domains. Data have been synthesised using Zweig and Shera's phenomenological model (1995) with a random scattering potential. Panel (a) shows the cochlear reflectance (solid line = real part; dotted line = imaginary part); (b) shows the normalised SFOAE resulting from the cochlear reflectance with the parameters, $g = 2$; $r = 1$ both assumed independent of frequency (solid line = real part; dotted line = imaginary part); (c) and (d) are the Fourier transforms of (a) and (b) respectively, where ϕ is the conjugate Fourier variable to η.

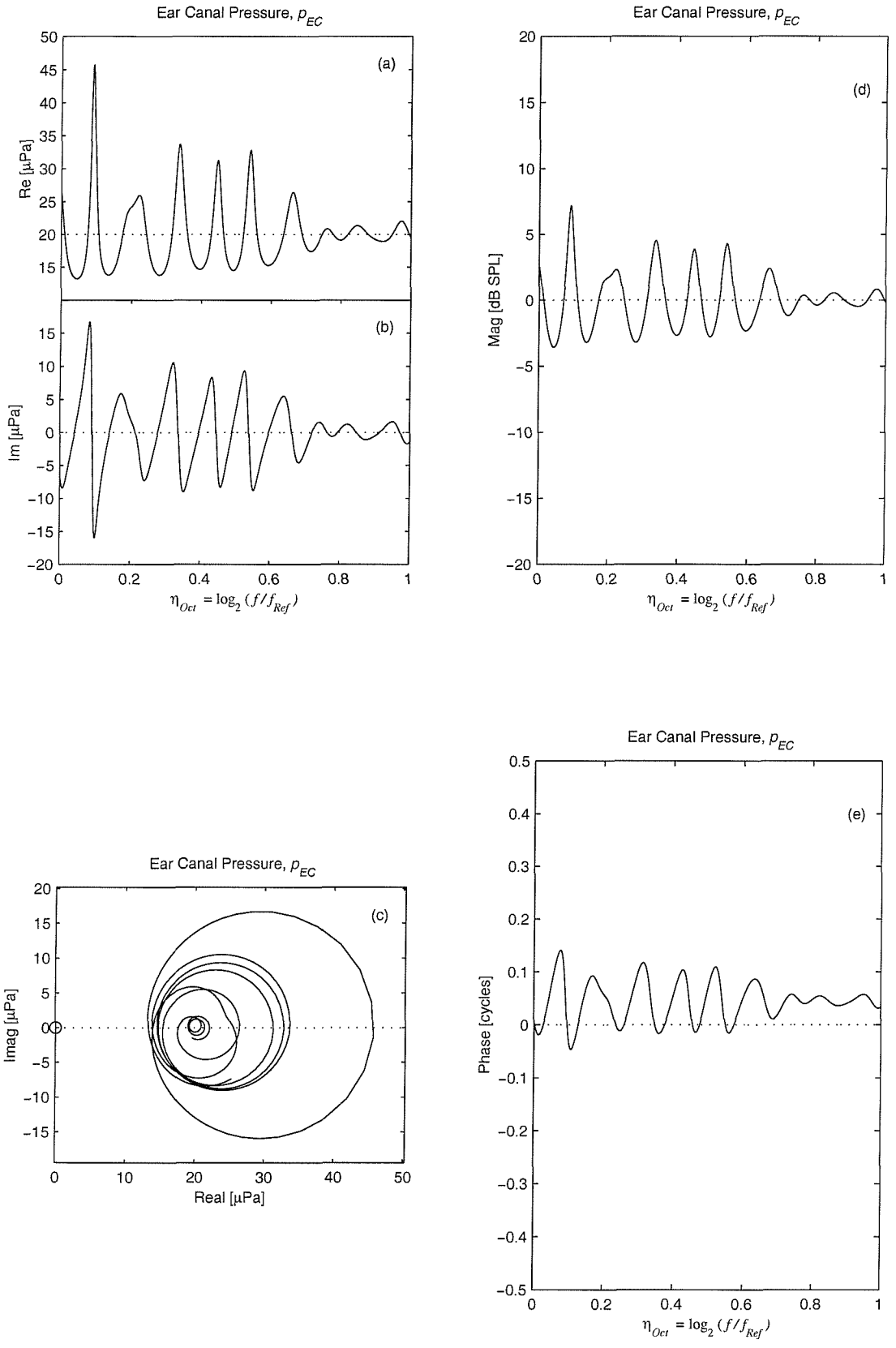


Fig 4.9. Representations of the complex ear canal pressure synthesised using Zweig and Shera's phenomenological model (1995) with a random scattering potential, and assuming that $g = 1$; $r = 1$; $p_{EC:R=0} =$ 20 μPa in equation [4.30]. Panels (a) and (b): the real and imaginary parts plotted against the η_{Oct}. Panel (c): real part plotted against the imaginary part. Panels (d) and (e): the magnitude and phase of the reflectance plotted against the η_{Oct}-variable. Solid line = ear canal pressure; dotted line = ear canal pressure for a reflectionless cochlea.

Fig. 4.7 shows alternative representations of p_Δ in the η–domain, for a case where multiple reflections are significant (r = 1, and the rms amplitude of R is 0.25). It shows similarities to the idealised OAEs seen in fig. 2.6, except for the random fluctuations in R. The effect of multiple reflections is to change the approximately sinusoidal ripple in Im{R} to a ripple in Im{ p_Δ } approaching a tangent function (fig. 4.7a and b). Also note from [4.30] the real and imaginary parts of p_Δ , like those of R, form a Hilbert transform pair (i.e., they are shifted by 90° relative to one another). Fig 4.7 (c) shows that the multiple reflections shift the locus of p_Δ to the right hand side of the complex plane. The size and direction of this shift depends on the magnitude and phase of rR. Fig 4.7 (d) and (e) show that multiple reflections cause ripples in the magnitude and phase of p_Δ that would otherwise be absent. These ripples have the same periodicity of about 1/15 as is seen in the real and imaginary parts. The average slope of the phase of p_Δ also equals the periodicity. From this it can be seen that the group delay at any given stimulus frequency, which is determined by the slope of the phase of p_Δ when plotted against linear frequency, will be related to the periodicity and the value of the stimulus frequency.

Generally $|rR|<1$ and therefore the binomial series expansion, [4.30], converges. However, it is postulated that sometimes, since R involves active amplification, it is possible that $|R|>1$ (Kemp, 1980; Zweig and Shera, 1995; Talmadge *et al.*, 1998). If $rR = 1$, such that a doubly reflected wave is neither reduced in amplitude, nor altered in phase, then the wave becomes self sustaining. As discussed in section 3.6, this phenomenon is thought to be the origin of SOAEs.

Fig 4.8 illustrates how the presence of multiple reflections appears in the ϕ-domain. The single pulse in $F\{R\}$ appears as a series of pulses of decreasing amplitude in $F\{p_\Delta\}$, centred at multiples of the fundamental ϕ-centre value (here equal to 15). These predicted forms of p_Δ and $F\{p_\Delta\}$ are important, because p_Δ (unlike R) is a directly measurable quantity.

The relationship between the actual ear canal pressure and the normalised SFOAE, can be found by rearranging [4.30], to give [4.31]:

$$p_{EC} = p_{EC:R=0}(1 + p_\Delta) \qquad [4.31]$$

This is illustrated in fig. 4.9 for a case where the zero-cochlear-reflectance ear canal pressure, $p_{EC:R=0}$ is independent of frequency. Unlike p_Δ , the ear canal pressure, p_{EC} shows ripples in both its magnitude and phase, even in the absence of multiple reflections. The

locus of $\boldsymbol{p}_{EC}$ in the complex plane is equal to that of $\boldsymbol{p}_{\Delta}$ (fig. 4.7 c), plus an offset of 1, and with an additional scale factor. See also the phasor diagram, fig 2.2, for comparison.

In the macromechanical models, the actual values of $\boldsymbol{g}$ and $\boldsymbol{r}$ depend on the front end model and on the impedance of the ear canal for a zero-reflectance cochlea. For the ideal front end model (section 4.6.5), $\boldsymbol{g} \approx 1$ and $\boldsymbol{r} \approx 0$, independent of stimulus frequency. Thus, in this case, basal reflections are negligible, and the SFOAE is approximately equal to the cochlear reflectance. For the high reflection front end model, $\boldsymbol{g} = 2$ (due to the pressure doubling) and $\boldsymbol{r} = 1$. For the Kringlebotn front end model (section 4.6.5), over the frequency range of interest (1.5–3 kHz), the value of $|\boldsymbol{g}|$ varies from –9 down to –26 dB, and of $|\boldsymbol{r}|$ from 0.4 up to 0.7. It is this fall off in the function, $\boldsymbol{g}$, that is responsible for the fall off in measured SFOAEs as the stimulus frequency increases above about 1.5 kHz.

4.10 Relating the Observed SFOAE Periodicity to Travelling Wave Shape

According to equations [4.26] and [4.30] the periodicity of the normalised SFOAE, $\boldsymbol{p}_{\Delta}$, is characterised by the location of the peak of the ϕ-spectrum, denoted by ϕ_C (cf. section 2.11). This in turn is related to the slope of the phase of the TW function near the TW peak, which is inversely proportional to the TW wavelength. This leads to the following relationships (Zweig and Shera, 1995):

$$\Psi \equiv \frac{\Delta f}{f_{GMF}} \approx \frac{1}{\varphi_C} \approx \frac{\hat{\lambda}}{2d} \approx 1/15 \text{ (humans)}$$

$$\Delta f \equiv f_2 - f_1 \qquad [4.32]$$

$$f_{GMF} \equiv \sqrt{f_1 f_2}$$

where f_1 and f_2 are adjacent peaks in the ripple in magnitude of $\boldsymbol{p}_{EC}$; ϕ_C is ϕ-centre value (i.e., the location of the peak of the ϕ-spectrum); $\hat{\lambda}$ is the TW wavelength near the TW peak; and d defines the place-frequency mapping (giving the distance along the BM over which the characteristic frequency changes by a factor of e). It has been estimated that $d \approx 7.2$ mm (Zweig and Shera, 1995), and thus, from the estimates of periodicity, Ψ, from measured SFOAEs, equation [4.32] gives an estimate of the peak TW wavelength of $\hat{\lambda} \approx 0.96$ mm. The importance of this is that any changes in the TW wavelength in the peak region will cause a corresponding change in the periodicity.

As well as the ϕ-centre value, ϕ_C , we are also interested in the ϕ-bandwidth. This is the width of the peak of $F\{\boldsymbol{p}_{\Delta}\}$ in the ϕ-domain, as seen in fig. 4.8d. This quantity is denoted by ϕ_{BW} and also shows up in the various η–domain representations in fig. 4.6 and

4.7. A large ϕ_{BW} shows up as large envelope modulations in the real and imaginary parts of $\boldsymbol{p}_{\Delta}$, and in the fluctuations seen in the magnitude and phase curves. The random nature of these fluctuations arises from the random scattering potential seen in fig. 4.4a, whilst the value of ϕ_{BW} is determined by the width of the peak in the $F\{\boldsymbol{T}^2\}$ curve. This in turn is inversely related to the width of the peak in of $\boldsymbol{T}^2$ in the η–domain. Thus [4.26] predicts that the sharper the TW peak the greater the ϕ-bandwidth, and a the less sharply defined the periodicity (i.e., the greater the modulation of the ripple pattern envelope). The physical explanation for this, according to Zweig and Shera (1995), is that the lower the 'Q' factor, the broader the TW peak, the more wavelengths it contains, the more coherent the scattering within the peak region, and thus the narrower the spread of periodicities in the SFOAE.

Thus the two quantities ϕ_C and ϕ_{BW} are related to two properties of the TW: the wavelength in the peak region, $\hat{\lambda}$, and the 'Q' factor of the envelope. Note, however, that these two properties are not wholly independent. According to Zweig and Shera (1995), the square of the TW function, $\boldsymbol{T}^2$, is minimum phase, which means that its magnitude and phase are not independent. Thus a reduction in the sharpness of the envelope (for example, by disturbing the cochlear amplifier) is accompanied by a reduction in the slope of the phase (corresponding to an increase in $\hat{\lambda}$). The significance of this is that changes in the travelling wave shape will result in changes in both ϕ_C and ϕ_{BW}.

Unlike the periodicity itself, this periodicity bandwidth has seldom been reported. One exception is the estimate given by Zweig and Shera (1995), obtained by averaging in the ϕ-domain across SFOAEs from several subjects, where they obtained $\phi_{BW} / \phi_C \approx 0.5$ (where ϕ_{BW} here denotes the full width of the peak, rather than the half-width used by Zweig and Shera, and which they denote $\Delta\varphi$).

4.11 Definition of Reflectance in Nonlinear Systems

For a linear system, the terms such as 'reflectance' or 'impedance' refer to complex quantities that are functions of frequency only. For a nonlinear system, the situation is much more complicated. In the case of SFOAEs, the terms reflectance and impedance become level dependent, and are defined as follows. Consider a single pure tone presented to the ear. The cochlear nonlinear response in the time domain can be separated into the primary or first order response (defined as the response component at the stimulus frequency), plus any distortion components, which are discarded. The reflectance can then

be defined in terms of the forward and backward components of the primary response. Frequency domain equations such as [4.30] can still be used, provided the two non-linear terms, $\boldsymbol{R}$ and $\boldsymbol{p}_{\Delta}$ are defined in terms of the primary values. (Note, however, that the interpretation of [4.30] as the sum of successive reflections can only be used with care). In the case of two tone suppression, $\boldsymbol{R}$ and $\boldsymbol{p}_{\Delta}$ (at the stimulus frequency) depend both on the level of the stimulus and of the suppressor. Also note that this nonlinear reflectance no longer gives a complete characterisation of the relationship between the reflected and incident TW.

5. Predicted SFOAEs from Macromechanical Cochlear Models

5.1 Objectives

The primary objective of the cochlear modelling is to obtain predictions of the behaviour of the SFOAE periodicity during induced changes in TW shape for two cochlear models: the first based on the Shera and Zweig's theory and the second based on Strube's theory. Changes in TW shape are to be induced by two methods: self-suppression and two-tone suppression. SFOAE periodicity is characterised by two signal parameters: the ϕ-centre value, ϕ_C, and the ϕ-bandwidth, ϕ_{BW}.

A secondary objective is to investigate the influence of various different features in the model in order both to better understand these features and to ensure that the model results are not overly sensitive to arbitrary choices of model formulation or model parameter values. To this end, the following are to be been studied:

1. the effect of a global change of cochlear amplifier gain,
2. the effect of different cochlear amplifier formulations (by comparing models containing the NK-1986 amplifier with those based on Z-1991 amplifier),
3. the effect of different spatial variations in the CP scattering impedance (in addition to the scattering impedances based on the theories of Strube and of Zweig and Shera),
4. the effect of any frequency dependency of the CP scattering impedance (by introducing a scattering impedance based on a spatially varying CP mass and stiffness in addition to that based on CP damping).
5. the effect of the size of BM inhomogeneities,
6. the effect of different middle ear models.

5.2 Specification of Model Variants and Input Stimuli

The models specified in section 4.6 have been implemented in Matlab 4.2 on a Pentium II PC. The Matlab script files defining the basic model are given in appendix II. Table 5.1 shows the model variants that were used, together with their stimuli. In all cases the stimulus frequency was swept between 1.5 and 3 kHz at 4 Hz intervals, whilst holding the OAE probe source level, Q_{Src} , constant, and whilst outputting the ear canal pressure at each frequency. This stimulus regime will be referred to as a 'frequency sweep'. The reason for choosing this frequency range is discussed in section 8.7.

5.2.1 Exploring the Basic Model Response

Models 1–10 are exploratory cases in which a single frequency sweep was performed with various different scattering impedances and cochlear amplifier impedances. In these cases, the ideal front end model (i.e., the transparent middle ear coupled to a non-reflecting OAE probe) was adopted. The purpose of these cases was to verify that the models gave similar predicted SFOAEs to those reported in previous publications, and to those from Zweig and Shera's phenomenological model. In addition both NK-1986 and Z-1991 cochlear amplifier formulations were used to check the sensitivity of the predicted SFOAEs to this basic feature of the models. Models 11 and 12 are as models 9 and 10, but with the high reflection front end model.

5.2.2 Modelling The Effect of Varying the Global Cochlear Amplifier Gain

Models 13-17 examine the effect on SFOAEs of altering the TW shape by globally reducing the cochlear amplifier gain. This is achieved by introducing a single attenuating factor into equations [4.12] and [4.13] for the cochlear amplifier impedance. For each value of the attenuating factor, a frequency sweep was obtained. The front end model based on Kringlebotn's model of the middle ear was used. Models 14 and 17 were chosen to test Zweig and Shera's theory with a random scattering impedance. Model 15 has a spatially periodic scattering impedance, as suggested by Strube (1989). Models 13 and 16 have a zero scattering impedance, and are included to define the reflectionless ear canal impedance, required to define the SFOAE pressure in equation [4.30]. As is shown later, this could also be achieved by eliminating the cochlear amplifier, rather than the scattering sites. For models 14 and 17, where the scattering impedance is random, each frequency sweep has been repeated n times, each time taking a different sequence of random numbers used to define the scattering impedance. Averaging was then performed across the resulting ensemble of SFOAE frequency sweeps. Typically, $n = 32$. The details of this averaging are discussed in section 5.5.

5.2.3 Modelling Self-suppression and Two-tone Suppression

Models 18 - 21 used Kanis and de Boer's quasilinear method to investigate the effect of self-suppression and two-tone suppression on SFOAEs. For the self-suppression simulation, frequency sweeps were obtained at various different levels of a stimulus tone.

This single tone will be referred to as the 'probe' tone (for compatibility of terminology with the two-tone suppression case). In the self-suppression simulations, the nominal probe levels ranged from 20 to 80 dB SPL.

For the two-tone suppression simulation, frequency sweeps were obtained whilst two tones are presented simultaneously, the first called the 'probe' tone, and the second the 'suppressor' tone. For each sweep, the levels of both the probe and suppressor tones were held constant. The degree of two-tone suppression is highly sensitive to the ratio of the suppressor frequency to the probe frequency. The suppressor tone is referred to as a 'low-side' suppressor when the ratio is less than 1, and a 'high-side' suppressor when it is greater than 1. In this thesis, only high-side suppression is considered. This is because models predict that the changes in TW shape induced by a low-side suppressor are similar to those induced by self-suppression (Kanis and de Boer, 1994). In high-side suppression, the models predict that the TW shape differs significantly from that in the self-suppression case (for a given TW peak amplitude) as will be further discussed in section 7.3.2. In the models used here, it was found that increasing the suppressor to probe frequency ratio at constant stimulus levels caused a increase in the difference between the shape of the two-tone suppressed TW and that of the self-suppressed TW. However, it also caused a reduction in the degree of suppression as measured by the magnitude of peak of the TW. As a compromise between the desire for a significantly different TW shape from the self-suppression case and the desire for a significant degree of suppression, a frequency ratio of 1.1 was chosen for the simulations reported here. As will be discussed in part III, it differs from the value of 1.3 used in experiments.

In two-tone suppression, the SFOAE values are obtained at the frequency of the probe tone. Frequency sweeps were obtained for nominal suppressor levels ranging from 30 to 80 dB SPL, and with a probe level of 45 dB SPL throughout. The suppressor to probe frequency ratio was held constant at 1.1 throughout. To test Shera and Zweig's theory, first a simulation with zero scattering impedance was performed to define the reflectionless ear canal pressure, $\boldsymbol{p}_{EC:R=0}$. Then a random scattering impedance was used to obtain the ear canal pressure, $\boldsymbol{p}_{EC}$, and thence the SFOAE pressure $\boldsymbol{p}_{SF}$. Owing to the computational time required to execute each frequency sweep, only two random number sequences were used.

5.2.4 General Points on the Models

In the above models, the size of the scattering impedance is controlled by the scale factor, a_{Sc} , given in equations [4.14] - [4.19]. This was set such that significant SFOAEs were generated, but without causing the model to become unstable. Typically the peak value of the normalised SFOAE was kept between 0.05 and 0.5. This is discussed further in section 5.8.

For all models the accuracy of the numerical method is determined by the spatial discretization. In all these models, at least 1024 equispaced points along the CP were used. This number was arrived at after checking the sensitivity of the results of several models to increases in the number of points.

Note also that in the forthcoming discussions, the following shorthand terms are used to describe the models.

'a passive model'	a model with zero cochlear amplifier impedance
'an active model'	a model with non-zero cochlear amplifier impedance
'a smooth model'	a model with zero CP scattering impedance
'a scattering model'	a model with non-zero CP scattering impedance
'a random scattering model'	a model with a spatially random CP scattering impedance as suggested in Shera and Zweig (1993b)
'a periodic scattering model'	a model with a spatially periodic CP scattering impedance as suggested in Strube (1989)

Table 5.1: Cochlear model variants used for SFOAE simulations

Model No.	Linear or Nonlinear	Cochlear Amplifier	CP Scattering Impedance	Front end model	Stimulus or run type
1	Linear	Passive	Zero	Ideal front end	Simple freq. sweep
2	"	Passive	Random	"	"
3	"	NK-1986	Zero	"	"
4	"	NK-1986	Step	"	"
5	"	NK-1986	Random	"	"
6	"	NK-1986	Alt. Random	"	"
7	"	NK-1986	Periodic	"	"
8	"	NK-1986	NB-random	"	"
9	"	Z-1991	Zero	"	"
10	"	Z-1991	Random	"	"
11	"	Z-1991	Zero	High reflect.	"
12	"	Z-1991	Random	"	"
13	"	NK-1986	Zero	Kringlebotn	Vary coch. amp. gain
14	"	NK-1986	Random	"	"
15	"	NK-1986	Periodic	"	"
16	"	Z-1991	Zero	"	"
17	"	Z-1991	Random	"	"
18	Nonlinear	NK-1986	Zero	"	Self-suppression
19	"	NK-1986	Random	"	"
20	"	NK-1986	Zero	"	Two-tone suppression
21	"	NK-1986	Random	"	"

Details of the table entries are given overleaf.

Details of Entries in Table 5.1

Column	Table entry	Details
2	Nonlinear	Quasilinear method (Kanis and de Boer, 1993b)
3	Passive [1]	No cochlear amplifier impedance. Passive model as KdB-1994 passive model Eqn. [4.10]
"	NK-1986 [2]	Cochlear amplifier as Neely and Kim (1986); Kanis and de Boer (1993b). Eqns [4.10] & [4.12]
"	Z-1991 [3]	Cochlear amplifier as in Zweig (1991); Talmadge *et al.* (1998). Eqns [4.11] & [4.13]
4	Zero	$\mathbf{Z}_{Sc} = 0$ for all x and f
"	Step	Step function. Eqn [4.18]
"	Periodic	Periodic scattering impedance as in Strube (1989). Eqn [4.16]
"	Random	Broad band random scattering impedance based on Zweig and Shera (1995) Eqn [4.15]
"	Alt. Random	Alternative random scattering impedance based on Zweig and Shera (1995) Eqn [4.19]
"	NB-Random	Narrow band random scattering impedance. Eqn [4.17]
5	Ideal front end	The ideal front end model, section 4.6.5
"	High reflect.	High reflection front end model, section 4.6.5
"	Kringlebotn	Kringlebotn front end model. section 4.6.5

Notes:

(1): The passive model here is the KdB-1994 passive model, Eqn. [4.10]

(2): The NK-1986 cochlear amplifier is always used in conjunction with the KdB-1994 passive model, Eqn. [4.10]

(3): The Z-1991 cochlear amplifier is always used in conjunction with the T-1998 passive model, Eqn. [4.11]

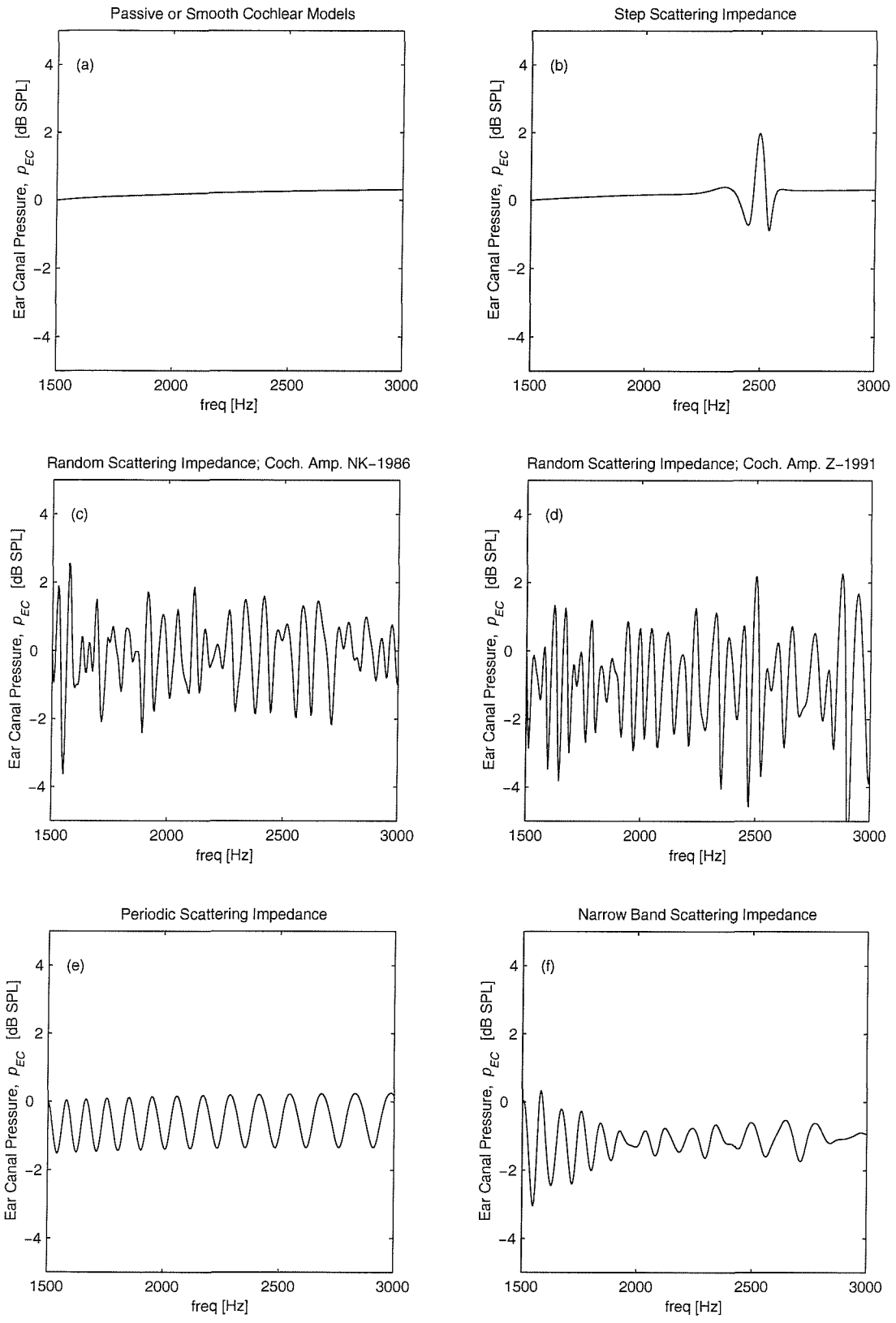


Fig 5.1 Ear canal pressure due to a constant volume velocity earphone simulated from various cochlear models. The level is set to give 0 dB SPL at 1.5 kHz in each case. Panel (a): three models which are almost indistinguishable: (i) passive and smooth (model 1); (ii) passive and rough (model 2); (iii) active and smooth (model 3). Panel (b): active model with step scattering impedance (model 4). Panel (c) active model (NK-1986 cochlear amplifier) with random scattering impedance (model 5). Panel (d): active model (Z-1991 cochlear amplifier) with random scattering impedance (model 10). Panel (e): active model with a periodic scattering impedance (model 7). Panel (f): active model with a narrow band random scattering impedance (model 8).

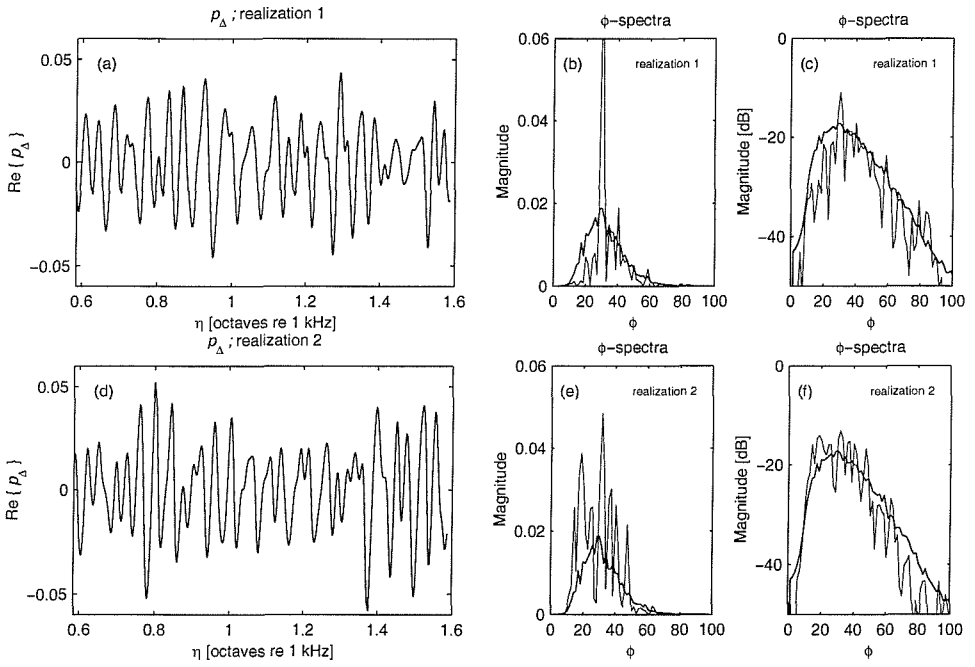


Fig 5.2 Comparison of two different realizations of the SFOAE simulated by a random scattering cochlear model (model 5). Results of normalised SFOAE, $\boldsymbol{p}_\Delta$, were obtained for 128 realizations of the scattering impedance. Two different realizations of the results are shown here, and compared with an ensemble average over the 128 realizations. Panel (a) Realization 1: $\boldsymbol{p}_\Delta$ against η_{Oct} . Panel (b) magnitude of the raw ϕ-spectrum for realization 1 (thin line) and the ensemble averaged ϕ-spectrum (thick line) plotted on a linear scale. Panel (c) as (b), but plotted on a dB scale. Panels (d)-(f) as panels (a) – (c), but for realization 2.

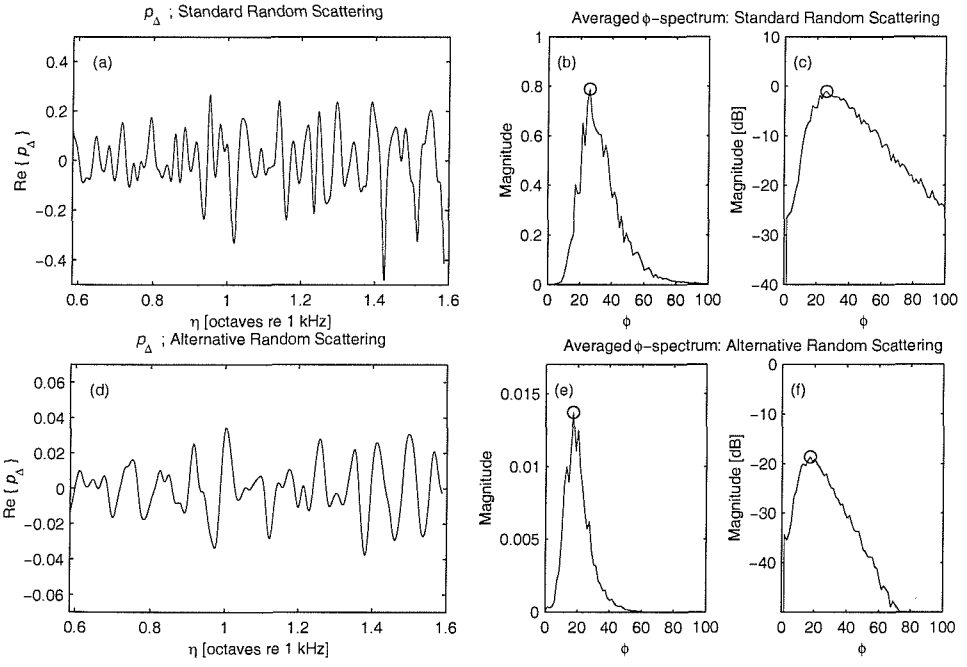


Fig 5.3 The effect on SFOAE periodicity of altering the formulation of the random scattering impedance in the cochlear model (models 5 and 6). Model 5 has the standard random scattering impedance, arising from the spatial variation in the CP damping. Model 6 has the alternative random scattering impedance, arising from the spatial variation in the total CP impedance. The predicted normalised SFOAE spectrum, $\boldsymbol{p}_\Delta$, was obtained for 32 realizations of each model. From these, the average ensemble averaged ϕ-spectrum was calculated. Panel (a): model 5, single realization of $\boldsymbol{p}_\Delta$ against η_{Oct} . Panels (b) and (c): model 5, magnitude of the ensemble averaged ϕ-spectrum plotted with linear and dB scales respectively. Panels (d)-(f) are the corresponding plots for model 6. Symbol o indicates the peak of the curve.

5.3 Discussion of Fundamental Model Results

Some fundamental results are shown in fig. 5.1. This plots the magnitude spectrum of the ear canal pressure for a single frequency sweep. Fig. 5.1a illustrates the expected result that no significant SFOAEs are generated in the absence of either the cochlear amplifier or of reflection sites. In this panel, the results from three models are overlaid and indistinguishable: the first is 'smooth and passive', the second is 'scattering and passive' and the third is 'smooth and active' where 'smooth' means the CP scattering impedance is zero; 'scattering' means the CP scattering impedance is non-zero; 'passive' means the cochlear amplifier impedance is zero; and 'active' means the cochlear amplifier impedance is non-zero (model numbers 1, 2 and 3, table 5.1). Fig. 5.1b shows that a simple step in the scattering impedance, which gives rise to a single point reflection site, generates significant reflections over a very limited frequency range. Fig. 5.1c and d show that the random scattering impedance together with the inclusion of either of the two cochlear amplifiers (models 5 and 10) give rise to a ripple pattern in which the peak-to-peak spacing increases with frequency, as predicted by Shera and Zweig's model. This result is insensitive to the choice of cochlear amplifier formulation.

Fig 5.1e confirms Strube's result (1989) that a periodic scattering impedance (together with a cochlear amplifier) gives rise to the basic periodicity in the SFOAE (model 7). However, these results are unrealistically regular. More realistic SFOAEs are shown in fig. 5.1f, resulting from a modified periodic scattering impedance, in which the spatial frequency of the corrugations has a narrow distribution rather a single value (model 8). This model can be thought of as lying somewhere between Strube's proposed periodic scattering impedance and Shera and Zweig's broad band random scattering impedance. Recall from section 4.8 that the essential difference between the two theories is that in Strube's theory, the periodicity is determined predominantly by the spatial periodicity in the CP scattering impedance, whereas in Shera and Zweig's model, it arises from the form of the TW function.

Results from model with the alternative random CP scattering impedance, based on spatial variations of the entire CP impedance, rather than of the damping (model 6) are not illustrated in fig 5.1, but are discussed in a later section.

5.4 Calculation of the Cochlear Model SFOAE Spectrum

The SFOAE spectrum for any active scattering model is calculated from the difference in the ear canal pressure obtained from two runs of the model: the first from the active scattering model, and the second from a smooth version of the same model, with identical stimulus conditions. This is given in equation [5.1] (as in equation [4.30]):

$$p_{SF} = p_{EC} - p_{EC:R=0}$$
$$p_{\Delta} = \frac{p_{SF}}{p_{EC:R=0}} \qquad [5.1]$$

where p_{EC} is the ear canal pressure for the active scattering cochlear model, and $p_{EC:R=0}$ is the reflectionless ear canal pressure defined here as the ear canal pressure for the smooth version of the cochlear model (i.e., one with the CP scattering impedance set to zero). Here p_{SF} is termed the SFOAE pressure, and p_{Δ} the normalised SFOAE. All quantities are complex functions of frequency. Thus, for example, the SFOAE for a random scattering impedance with the NK-1986 cochlear amplifier (model 5) is obtained from equation [6.1] by taking p_{EC} from model 5 and $p_{EC:R=0}$ from model 3. Similarly, the SFOAE for a random scattering impedance with the Z-1991 cochlear amplifier (model 10) is obtained by taking p_{EC} from model 10 and $p_{EC:R=0}$ from model 9.

The definition of the reflectionless ear canal pressure, $p_{EC:R=0}$, as that arising from the smooth cochlea requires further justification. The results in fig. 5.1a suggest that $p_{EC:R=0}$ can be estimated from either a passive cochlear model or from the smooth cochlea. However, in certain circumstances, when using Talmadge's model formulation (T-1998 with the Z-1991 amplifier) it was found that a slight discrepancy arose between the ear canal pressure predicted from an active-smooth model and that from a passive-smooth model (i.e., in these cases the curves corresponding to fig. 5.1a are not indistinguishable, but are slightly offset from one another). Recently Talmadge *et al.* (2000) have presented an analysis of such 'nonlinear SFOAE' components (i.e., components in the ear canal pressure arising not from the scattering impedance, but from nonlinear effects). According to Talmadge *et al.* (2000), these nonlinear SFOAEs are (i) highly sensitive to the formulation of the cochlear model, (ii) are of very low periodicity, especially above 1 kHz, and (iii) generally much lower in amplitude than SFOAEs arising from the scattering impedance. However, they may become influential at higher stimulus levels where nonlinear effects are larger and where SFOAEs due to the scattering impedance are smaller. The decision in this thesis to define the reflectionless ear canal pressure, $p_{EC:R=0}$, in terms of the smooth cochlea rather than the passive cochlea eliminates all such

'nonlinear' SFOAE components from the model predictions, leaving only SFOAEs arising from the scattering impedance. The advantage of this is that any changes induced in the predicted SFOAE spectra can be attributed purely to changes in the TW shape, thereby greatly simplifying the interpretation of the results. For a further discussion of the possible influence of nonlinear SFOAEs on the experimental results, see section 9.7.1

In the case of the cochlear models with random scattering impedance, the SFOAE spectrum depends not just on the model and its parameters, but also on the particular random number sequence used. In order to extract the underlying information from the SFOAE spectrum, a stochastic description is required, as explained in the next section.

5.5 Stochastic Description of the SFOAE Periodicity

The terminology in this section is taken from the theory of random processes. The problem addressed here is similar to that of calculating the power spectrum of a random process.

The results from a model with a random scattering impedance presented above depended on the exact form of the spatial variation in scattering impedance which resulted from the output of a random number generator. Re-running the same model with a different set of random numbers gives a different SFOAE function. What is desired are quantities to characterise the SFOAE signal that are dependent only on the underlying model formulation (which must include a statistical description of the random number generator), rather than on the particular sequence of random numbers. Such quantities are termed 'stochastic', because they pertain to an underlying probabilistic process, rather than to a single instance (or 'realization') of the process. As a simple example, a complete stochastic description of a gaussian white noise signal is given by specifying a single standard deviation, and by further specifying that the signal at every point in time is independent of that at every other point. A realization of this process is simply one actual white noise time history.

To obtain the stochastic quantities, models 5 and 10 were run n times with different realizations of the scattering impedance to give an ensemble of n realizations of normalized SFOAE frequency functions, $\boldsymbol{p}_\Delta(f)$. According to equation [4.30], $\boldsymbol{p}_\Delta$ depends on the reflectance, $\boldsymbol{R}$, and from equation [4.26], $\boldsymbol{R}$ looks like a narrow band random signal when plotted against the logarithmic frequency variable, η. (If equation [4.26] holds, and if the scattering potential is gaussian as it is in the model used here, then $\boldsymbol{R}$ is also gaussian.) Recall that a full stochastic description of the gaussian narrow band process is given by its

power spectral density, which can be estimated by averaging the power of the Fourier transform across an ensemble of realizations of the process (e.g., Newland, 1984). Although $\boldsymbol{p}_\Delta$ is not strictly gaussian narrow band (unless $\boldsymbol{r} = 0$) it can still be usefully described by a power spectrum. The method of calculating the power spectrum of $\boldsymbol{p}_\Delta$ follows standard spectral analysis methods (Newland, 1984). First $\boldsymbol{p}_\Delta$ must be in the form of a sequence at equal η intervals. All of the n realizations of $\boldsymbol{p}_\Delta$ are then Fourier transformed from the η-domain to the ϕ-domain. The n Fourier transforms are then multiplied by their complex conjugates, giving the square of the modulus of the transforms. This gives n functions, each of which will be termed a 'SFOAE raw ϕ-spectrum', where 'raw' indicates that no averaging has been performed. Averaging this across this ensemble of n raw ϕ-spectra yields an improved estimate of the underlying stochastic function. The underlying function will be termed the 'true ϕ-spectrum' and the estimate will be called the 'SFOAE average ϕ-spectrum'. Unlike the raw ϕ-spectrum, the true ϕ-spectrum is dependent only on the deterministic quantities in the model, and on the statistics of the random number generator. Figs 5.2a and 5.2b, show two realizations of SFOAE frequency functions obtained from an active random scattering models (model 5) using equation [5.1]. Their two corresponding raw ϕ-spectra are shown in figs 5.2c and 5.2d, together with an estimate of the true ϕ-spectrum, obtained from averaging across 128 realizations. The large variability between the two raw ϕ-spectra shows the difficulty in trying to estimate the true ϕ-spectrum, based on a single realization. In this project, the ϕ-spectrum will be characterised by two main quantities called the ϕ-centre value, denoted ϕ_C and the ϕ-bandwidth, denoted ϕ_{BW}. The ϕ-bandwidth is roughly equivalent to the 3 dB bandwidth of the peak. A method of estimating these two quantities is discussed in section 6.4.

5.6 The Effect of Different Formulations of the Random Scattering Impedance

Fig. 5.3 shows the SFOAEs resulting from two formulations of the random scattering impedance. The first (model 5) has the standard random scattering impedance, obtained by a spatial variation of the CP passive damping. The second (model 6) has the alternative random scattering impedance, obtained by a spatial variation of the entire CP impedance. Although both show the expected narrow band ripple pattern, they differ in their values of periodicity. This is clearly seen in the ensemble average ϕ-spectra, where for model 5, $\phi_C \approx 26$ (corresponding to a periodicity of about 4%), whilst for model 6, $\phi_C \approx 17$ (corresponding to a periodicity of about 6%). Model 6 also gives a narrower ϕ-spectrum.

The reason for this is that the alternative scattering impedance is frequency dependent, thereby invalidating the simple spatial filtering result presented in [4.26]. This effect is mentioned by Zweig and Shera (1995, p. 2033).

Although the alternative scattering impedance gives a periodicity that is closer to measured values, it was decided to adopt the frequency independent random scattering impedance as the standard for this thesis. This has the advantage of greater simplicity since it ensures that [4.26] becomes approximately valid, and consequently that the average ϕ-spectrum is simply a scaled version of the Fourier transform of the function $\boldsymbol{T}^2(\Theta)$. This result has been verified by calculating $\boldsymbol{T}(\Theta)$ for model 5. (Strictly, since this model does not exhibit perfect scaling symmetry, $\boldsymbol{T}(\Theta)$ is not uniquely defined for model 5, but is instead a function of both x and ω independently. However, because the degree of scaling symmetry is still high, this discrepancy has been found to be small. Thus a useful estimate of $\boldsymbol{T}(\Theta)$ can be made by evaluating both $\boldsymbol{T}$ and Θ as functions of cochlear location, x, at one constant stimulus frequency, ω.)

All the results reported in this thesis regarding changes in periodicity with TW shape have also been verified for both the standard and the alternative scattering impedances. For brevity, only the standard results are illustrated. Investigations have also been carried out in which the a spatial variation in the CP mass, stiffness or characteristic frequency was introduced. These also show similar trends to those seen for the standard case. This demonstrates that in these alternative cases the ϕ-spectrum is still related to the TW shape, even though the relationship is not the simple one given in [4.26].

5.7 The Absolute Value of the Predicted SFOAE Periodicity

The ϕ-spectrum seen in fig. 5.2 for model 5 is bandpass (in the ϕ-variable) with its centre located at a value, ϕ_C, of about 26 cycles (equivalent to a periodicity of about 1/26 or 4%). This periodicity is roughly half the value of around 7% reported in the literature. There are many possible reasons for this discrepancy, but these have not been investigated in detail. The models are based on many simplifying approximations (e.g., the long wave approximation) and contain many parameters whose values are uncertain[30]. For example, as will be seen later, simply reducing the cochlear amplifier gain significantly reduces the periodicity. The detail of the formulation of the random scattering impedance also has a significant effect, as discussed in the previous section. For these reasons, this discrepancy

[30] This discrepancy has recently been acknowledged by Talmadge *et al.* (2000) who also suggest alternative model parameters.

in the absolute periodicity is not thought to be a problem for the models. What is important for this thesis is not the absolute value of the periodicity, but the way in which it changes with the TW shape. Note that a similar discrepancy is seen in model 10 with the Z-1991 cochlear amplifier and the standard scattering impedance.

5.8 The Effect of the Size of the Scattering Impedance

The size of the scattering impedance is controlled by the scale factor, a_{Sc} , given in equations [4.14] - [4.19]. It was found that, for low values of the scattering impedance, scaling up the scattering impedance simply caused a corresponding scaling up in the SFOAE frequency spectrum, without any additional changes in its shape. This is expected from Zweig and Shera's phenomenological model (1995), equation [4.26], where scaling up the scattering potential causes a scaling up of the cochlear reflectance, $\boldsymbol{R}$. The physical explanation for this is that the scattered TW is much smaller than the forward TW, and therefore the CP response to the forward wave is approximately independent of the scattering sites. There is, however, an upper limit to value of this scale factor. For very large values, equation [4.26] no longer holds, as the wave travelling back from one reflection site becomes further scattered by more basal reflection sites, leading to multiple reflections which alter the shape of the spectrum of the cochlear reflectance. These should not be confused with the multiple reflections arising from the stapes, which affect the SFOAE rather than the cochlear reflectance, and which are accounted for in equation [4.30]. In all the models, a_{Sc} was set below this upper limit.

For the models with the ideal front end, where multiple reflections are negligible, the size of the scattering impedance is of little interest (provided the upper limit discussed above is avoided). In these cases, a_{Sc} was set to give typical normalised SFOAEs between 0.05 and 0.3. For the models with reflecting front ends, the value of a_{Sc} is more important, since a simple scaling up of the cochlear reflectance, $\boldsymbol{R}$, causes a more complex change in the normalised SFOAE. More specifically, as the value of $\boldsymbol{rR}$ increases, the influence of multiple reflections increases. In these cases, a_{Sc} was set to give a reasonably high cochlear reflectance magnitude of between 0.2 and 0.5 (in the unsuppressed state), such that some multiple reflection occurs. Note that in models with the Kringlebotn front end, the normalised SFOAE is less than the cochlear reflectance, due to the middle ear transmission loss related to the function, $\boldsymbol{g}$, in equation [4.30]. It was ensured that in all cases the value of a_{Sc} was not so high that instability occurred, due to runaway multiple reflections. Thus in these cases, there are two upper limits on a_{Sc} : the first to prevent significant multiple

reflections occurring from the CP scattering sites, the second to prevent unstable multiple reflections involving the stapes. The second turns out to be the lower, and therefore the important limit in these models.

5.9 Changes in Average SFOAE ϕ-spectrum with Cochlear Amplifier Gain

Fig. 5.4 shows the effect of varying the cochlear amplifier gain on the SFOAE ϕ-spectra for an active random scattering model (model 14). Figs 5.4a and b show the raw ϕ-spectra for two realizations of the SFOAE, whilst fig. 5.4c shows the average ϕ-spectrum estimated from an ensemble average over 32 realizations. The location of the peak in each spectrum, which gives a measure of the periodicity, is also marked. The change in periodicity with cochlear amplifier gain is clearly seen in the average ϕ-spectrum, fig. 5.4c. It is also shows up in the raw ϕ-spectra of realization 2, but not in those of realization 1. This shows that the changes in periodicity that are sought do not necessarily show up clearly in the raw ϕ-spectrum. To overcome this problem, a parametric model of the SFOAE ϕ-spectrum has been developed. This is described in the next section. Figs 5.4d-f show the results of this parametric fit to the two realizations, and to the ensemble average. Further discussion of the changes of periodicity due to varying the cochlear amplifier gain, self-suppression and two-tone suppression is presented after the development of this parametric model has been explained.

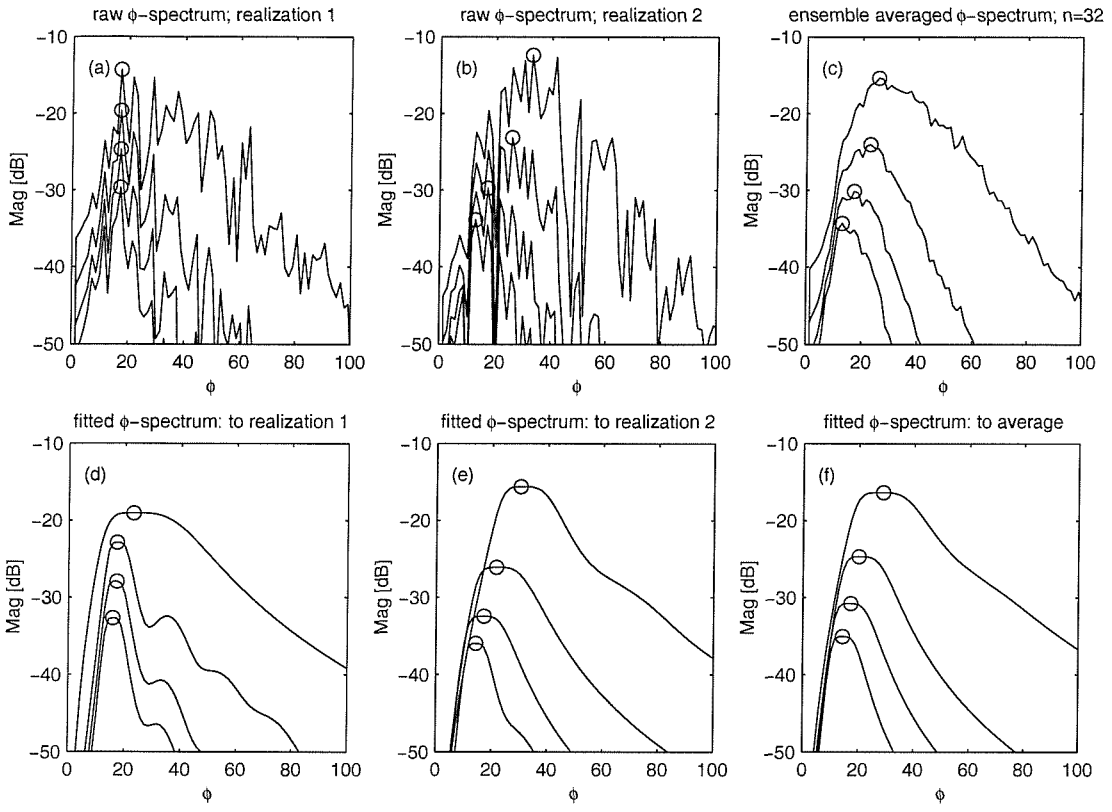


Fig 5.4 The effect of varying the cochlear amplifier gain on the SFOAE raw and average ϕ-spectra, simulated by a cochlear model with a random scattering impedance (model 14). The model has the NK-1986 cochlear amplifier attenuated by: 0, 0.4, 0.8, 1.2 and 1.6 dB. Panels (a) and (b) show raw ϕ-spectra resulting from two different realizations of the random scattering impedance. Panel (c) shows the average ϕ-spectra obtained from an ensemble of 32 realizations. Panels (d)-(f) show fitted ϕ-spectra obtained by applying the 4-parameter model to the data shown in panels (a)-(c) respectively. Symbol o indicates the peak of the curve.

6. A Parametric Model of SFOAE Frequency Spectra

6.1 Overview of Problem

The general question addressed in this section is how to characterise the SFOAE periodicity and thus allow comparisons to be made between predictions from the macromechanical models and experimental measurements. As has already been discussed, the Shera and Zweig's theory predicts changes in the SFOAE spectral periodicity with changes in the effective cochlear amplifier activity. In particular, the model predicts changes in terms of the ϕ-centre value, ϕ_C , and the ϕ-bandwidth, ϕ_{BW} , of the SFOAE average ϕ-spectrum. The problem highlighted in the previous section is that for the random scattering cochlear models, ϕ_C and ϕ_{BW} can only be directly calculated from the SFOAE average ϕ-spectrum, obtained by running a macromechanical model many times with different realizations of the random scattering impedance. This option is obviously not available with measured SFOAEs where only a single realization is available. Note that ensemble averaging across subjects does not solve the problem, as will be discussed in section 6.3.

In this section, various methods from the literature for estimating the SFOAE periodicity are reviewed. Also a description is given of two parametric models (called here the 3-parameter and 4-parameter models) which have been developed for analysing measured data.

6.2 Review of Parametric Spectral Analysis Techniques

To a first approximation, the SFOAE frequency spectrum, when viewed on a log frequency scale, can be seen as section of a stationary, bandpass random signal (equation [4.26]). The problem of estimating ϕ_C and ϕ_{BW} is similar to that of estimating the power spectral density of an unknown random process, based on a short section of a single realization.

One way of viewing any stationary random signal (usually a time series) is to imagine that the measured signal has arisen from an unknown white noise input signal passing through an unknown filter. In general, the purpose of spectral analysis is then to estimate the shape of the unknown filter, which then gives the power spectral density of the random process. (Our problem is a simpler one: to estimate the centre frequency and bandwidth of an unknown bandpass filter.)

The standard Fourier based approach to this problem is to use either segment averaging or spectral smoothing. In the segment averaging approach, the time sequence is split up into segments and the segments are then windowed, Fourier transformed, and power averaged. In the spectral smoothing approach the entire time sequence is windowed, Fourier transformed, and then smoothed in the frequency domain. However, these approaches performs poorly for applications where their is only a short sequence of data (Ables, 1978). There are several reasons for the poor performance of Fourier based method: they make unjustified assumptions about data that are not available (because the Fourier method implicitly assumes that the unavailable signal outside the window is either zero, or is a perfect periodic continuation of the available signal); and they impose distorting transformations on the available data (by applying an arbitrary tapering window). The effect of these assumptions is to give poor spectral resolution.

Parametric spectral analysis is an alternative approach in which the unknown filter is first characterised by a number of digital filter coefficients. The spectral analysis problem then becomes one of directly estimating these coefficients (Burg, 1978a, 1978b; Gutowski *et al.*, 1978). Unlike Fourier methods these methods attempt to obey Jaynes' principle of data reduction: "The result of any transformation imposed on the experimental data shall incorporate and be consistent with all relevant data and be maximally non-committal with regard to unavailable data" (Ables, 1978). Such methods can give much sharper spectral resolution than Fourier methods. The approach adopted in this thesis is loosely based on these parametric methods.

6.3 A Review of Periodicity Measures in Literature

In this section some of the shortcomings of the published methods of estimating SFOAE periodicity are discussed. These have already been introduced in section 2.12. The analysis of these periodicity measures has been performed using two methods of simulating SFOAE signals. The first generates SFOAE signals using cochlear models whilst the second generates simple bandpass random signals by passing white noise through a bandpass filter. In the latter method, the SFOAE ϕ-centre value corresponds to the centre frequency of the bandpass filter. In fact, both these methods reveal the same shortcomings in the published periodicity measures. Thus, although the SFOAEs generated from cochlear models will differ in some respects (i.e., in their full stochastic description) from simple bandpass signals discussed here, they show the same essential features which cause problems for these periodicity measures.

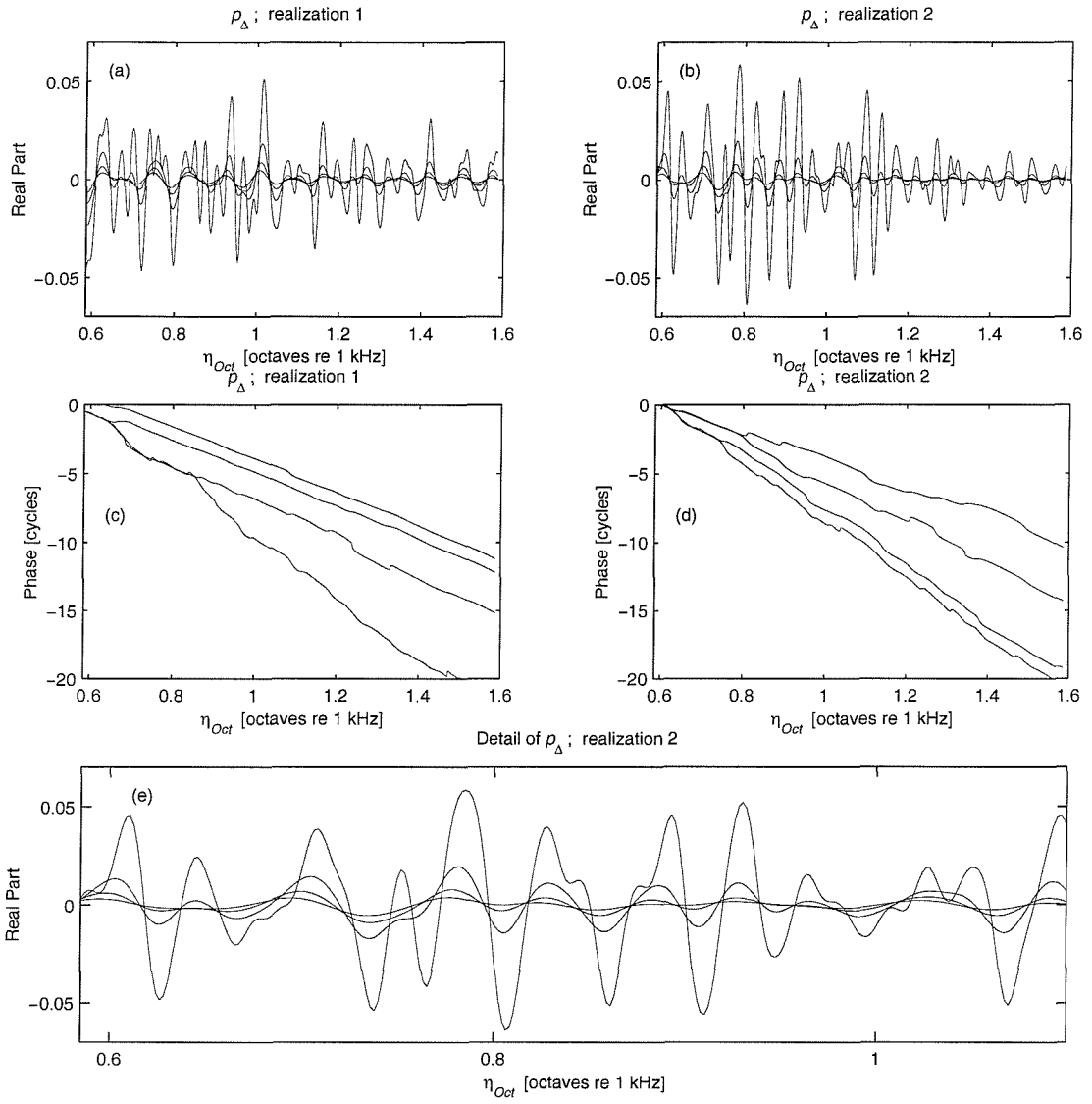


Fig 6.1 The effect of varying the cochlear amplifier gain on the SFOAE seen in different representations of the data. The SFOAE was simulated by a cochlear model with a random scattering impedance, and with NK-1986 cochlear amplifier attenuated by: 0, 0.4, 0.8, 1.2, 1.6 dB (model 14). Panels (a) and (b): two realizations of the normalised SFOAEs (Re$\{\boldsymbol{p}_\Delta\}$ against η_{Oct}). Panels (c) and (d): two realizations of the phase of normalised SFOAEs (arg$\{\boldsymbol{p}_\Delta\}$ against η_{Oct}). The slope of this gives a measure of the periodicity. Panel (e): detail of realization 2, showing the difficulty in clearly identifying peaks.

The published methods of estimating the periodicity are discussed with reference to fig. 6.1, which shows the effect on the SFOAEs of reducing the cochlear amplifier gain for the NK-1986 case (model 12). Two realizations are shown.

The first method is that of measuring peak-to-peak frequency intervals (Dallmayr, 1987; Zwicker and Schloth, 1984; Zwicker, 1990; Zwicker and Peisl, 1990; Lonsbury-Martin *et al.*, 1990). In this method, the turning points of the signal are first extracted, and then each peak-valley-peak sequence used to define one cycle. The difficulty with this method is that the signal comprises a broad range of large and small amplitude cycles. Some of these cycles are barely discernible (see for example, fig. 6.1f), and would easily be

lost in the presence of noise . This makes the method quite susceptible to contamination by noise. One way around this is to reject any cycles below some threshold of amplitude, but this introduces an arbitrary cut-off. A better method of estimating an average periodicity would be to apply some form of weighting based on the cycle amplitude. This is in effect what the next method, based on the Fourier transform, does.

In the second method the SFOAE frequency spectrum[31] is first mapped onto a log frequency scale (such as the η scale), and then transformed to the ϕ-domain using the Fourier transform (e.g., Lutman and Deeks, 1999). This method is similar to using the raw ϕ-spectrum directly as an estimate of the true ϕ-spectrum and has the problem that the raw ϕ-spectrum exhibits large peaks which arise purely from the random nature of the signal, rather than from the true underlying spectrum of the process. (When viewing the raw ϕ-spectrum as an estimator, its sampling distribution has a high standard deviation in relation to its mean; Bendat and Piersol, 1966). In fig 5.4a it can be seen that the peak of the raw ϕ-spectrum does not coincide with the peak of the true spectrum. Also, for this particular realization, as the cochlear amplifier gain is reduced, the location of raw ϕ-spectrum peak remains stubbornly unchanged. This is a phenomenon that can also easily be reproduced for bandpass noise obtained by simply passing a fixed white noise sequence through a bandpass filter several times, each time with a slightly different centre frequency.

The problem of the variability of the raw ϕ-spectrum could be approached in several ways. For example, as in standard spectral analysis, some smoothing in the ϕ-domain could be applied. However, this leads to a reduction in spectral resolution, as discussed in section 6.1. Zweig and Shera (1995) perform ensemble averaging of ϕ-spectra across subjects, but this method is not of use in studying changes in SFOAEs for a single subject. Also, it does not estimate the parameters of the random process of interest: the randomness in the SFOAE of an individual associated with the random scattering impedance alone. Instead, it includes the effect of intersubject variability of the entire cochlea (which would have to be characterised in a cochlear model by including a random distribution of cochlear parameters such as CP stiffness, or channel height).

A third method of estimating the periodicity is to calculate the group delay from the slope of the phase SFOAE frequency spectrum (Kemp and Chum, 1980a; Kemp and Brown; 1983). This is equivalent to measuring the instantaneous rate of rotation in the complex plane of the SFOAE vector as a function of frequency. To estimate a periodicity, the SFOAE phase can be plotted on a log frequency scale, and a best fit straight fitted to

[31] or some other signal exhibiting the SFOAE ripple pattern such as the ear canal sound pressure level.

the data. One problem with this method is that the amplitude of the signal is ignored when fitting the straight line to the phase data. This means that where the signal is weak (due to random fluctuations) the phase is poorly defined (especially if noise is present). Even with perfectly clean data, this method is insensitive to changes in periodicity. This can be seen in figs 6.1c and d, which show the SFOAE phase for the two realizations. In fig. 6.1c, three of the curves appear to run virtually parallel to each other.

As with the previous method, this phenomenon can be demonstrated easily using a bandpass time series. (Recall that here we are considering an SFOAE plotted against log frequency as a time series.) For the time series, the group delay method corresponds to estimating the centre frequency of the filter using the instantaneous frequency of the output signal. (Note that the SFOAE is a complex signal: for real signals instantaneous frequency is defined via the Hilbert transform; Randall, 1987). Instantaneous frequency is a useful measure for a single component signal (e.g., for demodulating an FM signal), but is less useful for a broader band signals in which several frequency components are present simultaneously. Although averaging the instantaneous frequency does give some measure of the 'average' frequency in the signal, it is a highly variable estimate, just like the location of the peak in the raw spectrum.

In summary, the published methods of estimating periodicity all have shortcomings which make the detection of small changes in periodicity quite difficult, and may explain why changes with level (for example) have not been more widely reported previously (section 2.11). In general, the published methods would work well for a very narrow band signal, but are progressively less accurate as the bandwidth is increased.

6.4 Description of a Parametric Model of SFOAE Signals

From equation [4.30], it can be seen that the predicted SFOAE η-function, $\boldsymbol{p}_{\Delta}(\eta)$ can be considered as a stationary bandpass random signal provided that three conditions are met: first if there is very little basal reflectance (i.e., $\boldsymbol{r}$ is negligible); second, if the middle ear transmission function, $\boldsymbol{g}$, does not vary significantly over the measured frequency range; and third, if the cochlear reflectance, $\boldsymbol{R}$, can itself be treated as a stationary bandpass random signal. That $\boldsymbol{R}$ can be treated as a stationary bandpass random signal is a prediction of the Zweig and Shera's theory, provided first that scaling symmetry roughly holds and second that the scattering potential is a stationary random function of place (as seen in the spatial filtering result in equation [4.26]).

Two parametric models of the SFOAE η-function have been developed. The first is a 3-parameter model which assumes that the function can be treated as a realization of a stationary bandpass random process. The second is a 4-parameter model which takes some account of multiple reflections due to the basal reflectance, $\boldsymbol{r}$. In both cases, any nonstationarity due to variation of $\boldsymbol{g}$ or $\boldsymbol{r}$ with frequency, or due to the spatial form of the scattering potential has been ignored. The justification for this is discussed in section 6.5.

The approach taken here does not attempt to adhere rigorously to Jaynes' principle (section 6.2) which is theoretically problematic. Instead a technique has been developed here that has been found to work when applied to results from macromechanical models. Furthermore, it is not suggested that this is an optimal technique - only that it is a useful one. The parametric methods described by Burg (1978a, 1978b) and Gutowski *et al.* (1978) are not directly applicable, because they estimate the coefficients of a general digital filter, rather than assuming a bandpass filter at the outset. The chosen method here does, however, borrow some elements from those methods.

6.5 The 3-Parameter Model of SFOAE

The 3-parameter model assumes that the η-SFOAE function can be treated as a realization of a stationary bandpass random process. Unlike general spectral estimation, we already have some idea of the nature of the filter (via the macromechanical model), and we are only looking to estimate two parameters that characterise the shape of the spectrum (i.e., ϕ_C and ϕ_{BW}). In this section, a parametric model is described which estimates these two parameters, and then calculates a third parameter, $p_{\Delta:\,\mathrm{RMS}}$, which is simply the RMS value of $\boldsymbol{p}_\Delta$ when considered as a function of η.

For this 3-parameter model, the first assumption is that the cochlear reflectance is a stationary bandpass signal in the η-domain, $\boldsymbol{R}(\eta)$. This follows from [4.26], with its assumptions that the scattering potential is a stationary white (or broad band) random process in the spatial domain, and that scaling symmetry holds.

The second assumption in the 3 parameter model is that multiple reflections are small, such that the basal reflectance, $\boldsymbol{r}$, in [4.30] can be taken as zero. It is then further assumed that the function, $\boldsymbol{g}$, in [4.30] is independent of frequency, and has a value denoted $\boldsymbol{g}_0$ such that:

$$\boldsymbol{p}_\Delta(\eta) = \boldsymbol{g}_0 \boldsymbol{R}(\eta) \qquad [6.1]$$

Thus the measured normalised SFOAE, $\boldsymbol{p}_\Delta$ is also a stationary bandpass η-function.

Then, inspection of the results of the macromechanical models shows that the shape of the SFOAE average ϕ-spectrum is well represented by a 2nd order Butterworth bandpass filter in the ϕ-domain. This is shown in fig. 6.2a-c for model 5, which has the NK-1986 cochlear amplifier, and in figs 6.2d-f, for model 10, which has the Z-1991 cochlear amplifier. This filter is specified by only two parameters: the upper and lower cut-off frequencies, or equivalently, the bandwidth and centre frequency. Therefore it was decided to define the ϕ-centre value and ϕ-bandwidth (ϕ_C and ϕ_{BW}) of the SFOAE data by the corresponding parameters for a 'best-fit' 2nd order Butterworth filter.

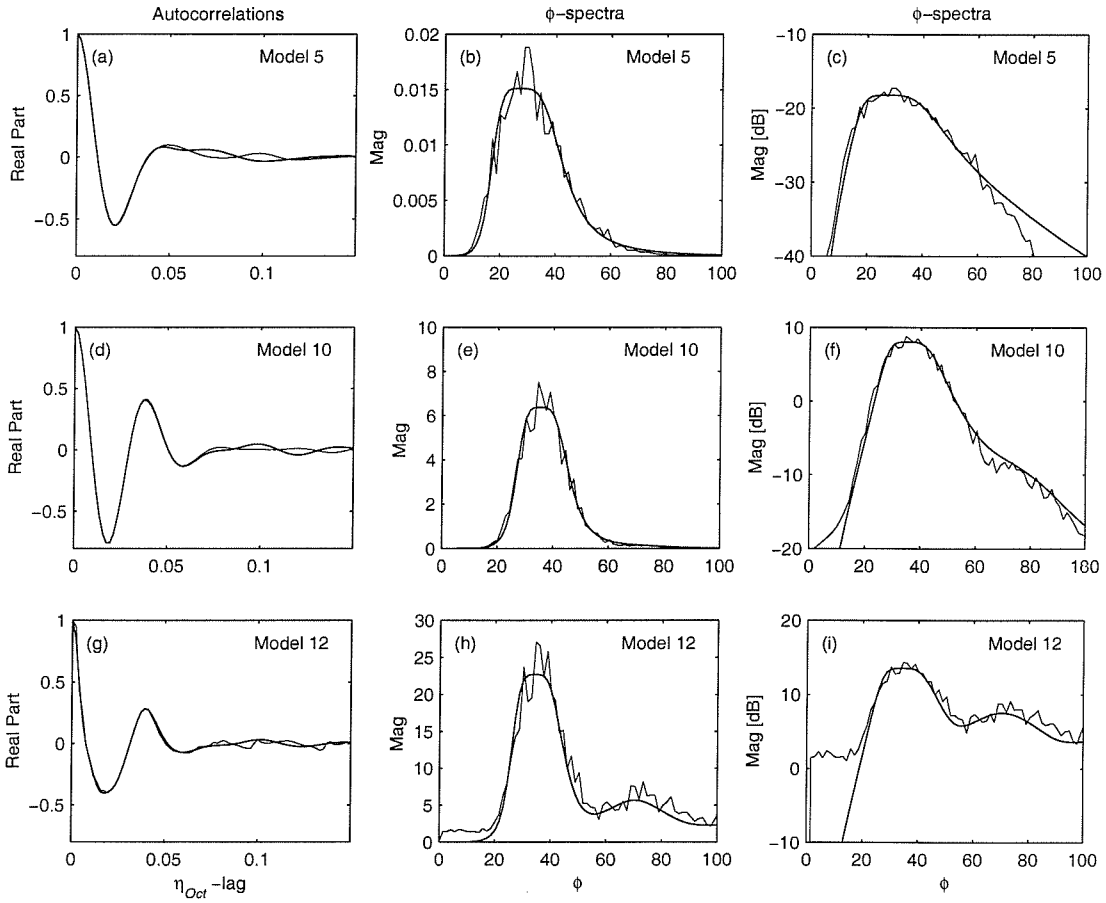


Fig 6.2 Comparison of the ensemble averaged SFOAE data with a parametric fit based on a 2nd order Butterworth filter. The average SFOAE data (thin lines) were obtained from 32 realizations from cochlear models with random scattering impedance. The fit (thick lines) is performed on the ensemble averaged normalised autocorrelation function using a 4-parameter model of the random process. The figure shows the results from three cochlear models: models 5, 10 and 12. Model 5 has the NK-1986 cochlear amplifier and an ideal front end model; model 10 has the Z-1991 cochlear amplifier and an ideal front end model; model 12 has the Z-1991 cochlear amplifier and the high reflecting front end model, which leads to significant mulitple reflections. Panels (a)-(c): model 5, normalised autocorrelation function, ϕ-spectrum (linear scale) and ϕ-spectrum (dB scale) respectively. Panels (d)-(e) corresponding plots for model 10. Panels (f)-(g) corresponding plots for model 12.

An iterative procedure has been adopted for fitting the ϕ_C and ϕ_{BW} values to the measured data. It turns out that there are advantages to fitting these parameters to an estimate of the autocorrelation function formed from the data, rather than to the raw ϕ-spectrum, as discussed in appendix III. The first step in the fitting procedure is to assume initial values of ϕ_C and ϕ_{BW} thereby defining a 2nd order Butterworth filter. Assuming a stationary white noise input to the this filter, the autocorrelation function of the output is then defined purely by ϕ_C and ϕ_{BW}. Thus treating the reflectance, $\boldsymbol{R}(\eta)$, as the output of the filter, and $\boldsymbol{h}(\eta)$ as the impulse response function of the filter, a standard result is that the autocorrelation function of the reflectance is related to the filter impulse response by:

$$\begin{aligned} &R(\eta) = h(\eta) \otimes w(\eta) \\ &\Rightarrow \mathsf{R}_{RR}(\eta') \propto \mathsf{R}_{hh}(\eta') \end{aligned} \qquad [6.2]$$

where $w(\eta)$ is a white noise signal; $\otimes$ denotes convolution; and where:

$$\begin{aligned} &\mathsf{R}_{RR}(\eta') \equiv E\left[\boldsymbol{R}^*(\eta)\boldsymbol{R}\ (\eta+\eta')\right] \\ &\mathsf{R}_{hh}(\eta') \equiv \int_{-\infty}^{\infty} \boldsymbol{h}^*(\eta)\boldsymbol{h}\ (\eta+\eta')d\eta \end{aligned} \qquad [6.3]$$

Here, $\mathsf{R}_{RR}(\eta')$ is the autocorrelation function of the reflectance, $\boldsymbol{R}(\eta)$; * denotes the complex conjugate; and E[] denotes the expectation operator. The assumption that $\boldsymbol{R}$ is stationary ensures that the autocorrelation function is independent of absolute values of η, and is instead a function of the a variable called here η-lag (denoted here by η') which defines the shift in η. The impulse response function, $\boldsymbol{h}(\eta)$, is a complex version of the standard 2nd order Butterworth filter impulse response, which has been chosen to ensure that $\boldsymbol{R}$ has the analytic properties discussed in section 2.13. (Recall that the Fourier transform of $\boldsymbol{R}$ is one-sided). To achieve this, the function , $\boldsymbol{h}(\eta)$, is defined as the analytic function corresponding to the wholly real impulse response function of the standard 2nd order Butterworth filter, denoted $h_{Re}(\eta)$. The definition of the analytic form is such that the transform of $\boldsymbol{h}(\eta)$ in the ϕ-domain is purely one-sided, and equal to the right hand side of the (two-sided, symmetric) transform of the corresponding real function, $h_{Re}(\eta)$.

It is useful to define a version of the autocorrelation function normalised with the variance the signal:

$$\mathsf{K}_{RR}(\eta') \equiv \mathsf{R}_{RR}(\eta')/\mathsf{R}_{RR}(0) \qquad [6.4]$$

where K_{RR} is the normalised autocorrelation, such that $K_{RR}(0)=1$. Note that the normalisation is performed using $R_{RR}(0)$ which is equal to the variance of $\boldsymbol{R}(\eta)$. From [6.2], the following relationship holds between normalised autocorrelation functions:

$$K_{RR}(\eta') = K_{hh}(\eta') \qquad [6.5]$$

and from [6.1] it also follows that:

$$K_{p_\Delta p_\Delta}(\eta') = K_{RR}(\eta') \qquad [6.6]$$

where K_{hh} and $K_{p_\Delta p_\Delta}$ are normalised autocorrelation functions for the normalised SFOAE, $\boldsymbol{p}_\Delta(\eta)$ and the impulse response function, $\boldsymbol{h}(\eta)$.

Therefore, given values of ϕ_C and ϕ_{BW}, the value of $K_{p_\Delta p_\Delta}$ can be calculated from [6.3], [6.4], [6.5] and [6.6]. An estimate of $K_{p_\Delta p_\Delta}$ can also be calculated from the actual measured η-series, $\boldsymbol{p}_\Delta(\eta)$, using the following estimator:

$$\tilde{R}_{p_\Delta p_\Delta}(\eta') \equiv \frac{1}{\eta_2 - \eta_1} \int_{\eta_1}^{\eta_2 - |\eta'|} \boldsymbol{p}_\Delta^{\,*}(\eta)\boldsymbol{p}_\Delta(\eta+\eta')d\eta \qquad [6.7]$$

$$0 < |\eta'| < \eta_2 - \eta_1$$

$$\tilde{K}_{p_\Delta p_\Delta}(\eta') \equiv \tilde{R}_{p_\Delta p_\Delta}(\eta') / \tilde{R}_{p_\Delta p_\Delta}(0) \qquad [6.8]$$

where the diacritical mark ~ denotes an estimate.

The full iterative procedure is as follows. Initial values of ϕ_C and ϕ_{BW} are assumed, thereby defining the impulse response of a (real) 2nd order Butterworth filter, $h_{Re}(\eta)$. The (complex) analytic signal, $\boldsymbol{h}(\eta)$, corresponding to $h_{Re}(\eta)$ plus the Hilbert transform of $h_{Re}(\eta)$ is then calculated in the standard way (e.g., Randall, 1987). From this the normalised autocorrelation function of the filter impulse response is calculated from [6.3] and [6.4]. Then the fitted version of the normalised autocorrelation function of $\boldsymbol{p}_\Delta(\eta)$, $K_{p_\Delta p_\Delta}$, is calculated from [6.6]. This is then compared with the measured estimate of the normalised autocorrelation function of $\boldsymbol{p}_\Delta(\eta)$, $\tilde{K}_{p_\Delta p_\Delta}$, obtained from [6.7] and [6.8]. A mean square error is then calculated:

$$\varepsilon_{MS}^2 \equiv \frac{1}{\eta_{Trunc}} \int_0^{\eta_{Trunc}} \left(\tilde{K}_{p_\Delta p_\Delta}(\eta') - K_{p_\Delta p_\Delta}(\eta')\right)^2 d\eta' \qquad [6.9]$$

that quantifies the poorness of fit between the fitted version and the measured estimates of the normalised autocorrelation function. The values of ϕ_C and ϕ_{BW} are varied iteratively to minimise this error. This error quantity is calculated over a restricted range of η-lag values, η_{Trunc}. One reason for this restriction in η' is that $\tilde{K}_{p_\Delta p_\Delta}(\eta')$ contains a high

degree of redundancy, such that most of the information about the underlying random process is contained in the early part of the signal. A second reason for this is that the estimates, $\tilde{K}_{p_\Delta p_\Delta}(\eta')$ become less reliable as η' approaches its maximum possible value, given by $\eta_2-\eta_1$ in [6.7]. These issues, together with the choice of η_{Trunc} are discussed in appendix III. The issues of convergence and of bias error are addressed in section 6.8.

The final parameter in the 3-parameter model is $p_{\Delta:\,RMS}$, defined by:

$$p^2_{\Delta:RMS} = \tilde{R}_{p_\Delta p_\Delta}(0) \equiv \frac{1}{\eta_2-\eta_1}\int_{\eta_1}^{\eta_2} \boldsymbol{p}_\Delta^{\,*}(\eta)\boldsymbol{p}_\Delta(\eta)d\eta \qquad [6.10]$$

and is referred to as 'the RMS normalised SFOAE'. When in decibel form, this will be referred to as the 'RMS normalised SFOAE level', denoted by $L_{\Delta:RMS}$.

This method of obtaining estimates of ϕ_C and ϕ_{BW} by characterising the SFOAE signal with the 3-parameter model will be referred to as the 3-parameter model.

6.6 The 4-Parameter Model

The 3-parameter model is based on the assumption that the SFOAE η-series can be considered as a stationary bandpass random signal. One problem with this model can be seen in figs 6.2g-i for model 12, where there is significant multiple reflection in the cochlea. In this case, the ϕ-spectrum is not purely bandpass, but shows a series of additional bandpass lobes (see also figs 2.6 and 4.8). This arises when the product $\boldsymbol{rR}$ becomes significant relative to 1. If the second lobe is significant relative to the main lobe, then the 3-parameter model will interpret multiple reflections incorrectly as spuriously high values of ϕ_C and ϕ_{BW}. To partially account for this affect, an additional parameter, denoted α, has been introduced. This gives some measure of the magnitude of $\boldsymbol{rR}$ averaged across the measured η range, and hence gives an indication of the strength of multiple reflections.

The 4-parameter model considers the case where the functions $\boldsymbol{g}$ and $\boldsymbol{r}$ in equation [4.30] do not very greatly over the signal, and can therefore be replaced by the constants $\boldsymbol{g_0}$ and $\boldsymbol{r_0}$ such that:

$$\boldsymbol{p}_\Delta = \frac{\boldsymbol{g_0 R}}{1-\boldsymbol{r_0 R}} \qquad [6.11]$$

It can then be shown that, if $\boldsymbol{R}$ is a gaussian stationary random process, then the following relationship holds:

$$\mathsf{R}_{p_\Delta p_\Delta} = g_0^2 \mathsf{R}_{RR}\left[1 + 2r_0^2 \mathsf{R}_{RR} + 6r_0^4 \mathsf{R}_{RR}^2 + \ldots\right] \qquad [6.12]$$

where $\mathsf{R}_{p_\Delta p_\Delta}$ is the autocorrelation function of $\boldsymbol{p}_\Delta(\eta)$, R_{RR} is the autocorrelation function of $\boldsymbol{R}$ defined in [6.2], g_0 is the magnitude of $\boldsymbol{g_0}$ and r_0 is the magnitude of $\boldsymbol{r_0}$. The derivation of this relationship is given in appendix III. For low values of r_0 the series expansion in [6.12] converges after only the first two or three terms.

As with the 3-parameter model, it is useful to normalise the autocorrelation functions in [6.12] by the variances of the signals giving:

$$\mathsf{K}_{p_\Delta p_\Delta}(\eta') = \frac{\beta^2}{\sigma^2_{p_\Delta p_\Delta}} \mathsf{K}_{\boldsymbol{RR}}(\eta')\left[1 + 2\alpha^2 \mathsf{K}_{\boldsymbol{RR}}(\eta') + 6\alpha^4 \mathsf{K}^2_{\boldsymbol{RR}}(\eta') + \ldots\right]$$

$$\alpha = r_0 \sigma_{RR} \qquad [6.13]$$

$$\beta = g_0 \sigma_{RR}$$

where, for an arbitrary signal, $x(\eta)$, $\mathsf{K}_{xx}(\eta')$ denotes the normalised autocorrelation function, such that $\mathsf{K}_{xx}(0) = 1$; and where σ_{xx} denotes the rms amplitude of $x(\eta)$.

The 4-parameter model fits the four parameters: ϕ_C, ϕ_{BW}, α and β required to define the right hand side of equation [6.13] to the measured data. The iterative fitting procedure is then as follows. First, the measured value of $p_{\Delta:\,\mathrm{RMS}}$ is calculated from [6.10]. This is used as an estimate of $\sigma_{p_\Delta p_\Delta}$. Then initial guesses are made of the three parameters: ϕ_C, ϕ_{BW} and α. From the values of α and of $p_{\Delta:\,\mathrm{RMS}}$, an estimate of the value of β can be calculated using [6.13], evaluated at $\eta' = 0$:

$$1 \approx \frac{\beta^2}{\sigma^2_{p_\Delta p_\Delta}}\left[1 + 2\alpha^2 + 6\alpha^4\right]$$

therefore [6.14]

$$\beta^2 \approx \frac{p^2_{\Delta:RMS}}{1 + 2\alpha^2 + 6\alpha^4}$$

From the values of ϕ_C and ϕ_{BW}, the fitted version of the normalised autocorrelation function of the reflectance, $\mathsf{K}_{RR}(\eta')$ can be calculated as for the 3-parameter model, using equations [6.3], [6.4] and [6.5]. The fitted version of the normalised autocorrelation function of the normalised SFOAE, $\mathsf{K}_{p_\Delta p_\Delta}$, can then be calculated using [6.13]. This is then compared with the measured estimate of the normalised autocorrelation function of $\boldsymbol{p}_\Delta(\eta)$, $\tilde{\mathsf{K}}_{p_\Delta p_\Delta}$, obtained from [6.7] and [6.8]. As for the 3-parameter model, a mean

square error term is calculated, using [6.9]. Iteration over the three parameters, ϕ_C , ϕ_{BW} and α continues until this error is minimised. All four parameters are now defined. The SFOAE ϕ-spectrum which can be calculated from the four parameters will be called the 'SFOAE fitted ϕ-spectrum'.

The result of applying the 4-parameter model to the output of cochlear models is shown in figs 6.2g-i, where the parametric fit is compared to the ensemble averaged SFOAE ϕ-spectrum for the model 12, where multiple reflections in the cochlea are significant. Note that if $\boldsymbol{r}_0 = 0$ then $\alpha = 0$, $\beta = p_{\Delta:\,\mathrm{RMS}}$ and the 3-parameter model is recovered.

This method of obtaining estimates of ϕ_C and ϕ_{BW} by characterising the SFOAE signal with the 4-parameter model will be referred to as the 4-parameter model. A Matlab procedure for returning the four parameters for any given SFOAE signal can be found in appendix II.

6.7 Validity of the 4-Parameter Model

With reference to equation [6.1], the assumptions underlying the 4-parameter model are that $\boldsymbol{g}$ and $\boldsymbol{r}$ are independent of frequency, and that $\boldsymbol{R}$ is a random signal arising from stationary, gaussian white noise passed through a 2^{nd} order Butterworth band filter, defined by ϕ_C and ϕ_{BW} . Although it is clear from our knowledge of the middle ear models that $\boldsymbol{g}$ and $\boldsymbol{r}$ are not independent of frequency, their variation with frequency is much slower than that of the SFOAE, $\boldsymbol{p}_\Delta$. Consequently it has been found in the models used here that, the 4-parameter model can still be usefully applied even when realistic middle ear models are included. The values of α and β that are then obtained relate to frequency averaged values of $|\boldsymbol{gR}|$ and $|\boldsymbol{rR}|$ rather than the values given in [6.13]. Similarly, the reflectance, $\boldsymbol{R}$, may not in reality satisfy all the above assumptions. This is because scaling symmetry is only approximate, the scattering potential may not be stationary and gaussian, and the spatial filtering equation [4.26] is only approximate.

6.8 The Performance of the 4-Parameter Model for Cochlear Model Data

It is useful to assess the performance of the 4-parameter model with regard to the variability of the estimates and to the sensitivity to noise. Quantifying the performance analytically is extremely difficult. Instead, the following numerical approach has been

taken. First data have been generated for which the optimal four parameters are already known. Then the four parameters are estimated from the 4-parameter model and the estimates compared to these optimal values. Three kinds of signal have been used to check the performance in this way. The first kind is an ideal synthetic signal which satisfies all the assumptions stated in section 6.4. This was achieved by passing a white noise signal through a known 2^{nd} order Butterworth bandpass filter, and then inserting the resulting bandpass signal into equation [6.1] to give an ideal output signal. This has been repeated many times in order to generate the sampling distribution of the estimators for the four parameters. The second kind of signal is an ideal synthetic signal plus additive random noise. This gives an indication of the signal-to-noise ratio (SNR) for which the model can be safely used. The third kind of signal is output generated by the macromechanical model.

These investigations show that for estimating ϕ_C , the performance of the 4-parameter model is much better than that of a crude Fourier based estimator, which returns the location of the peak of the raw ϕ-spectrum. (The variance of the parametric estimator is roughly a half that of the crude Fourier based estimator.) The results also suggest that the model performs successfully down to an SNR of about 0 dB. However, estimates of ϕ_{BW} and α are less reliable than those of ϕ_C and β. The results also show that for low values of α (<0.1) the 4-parameter model gives very similar results to the 3-parameter model. There is no significant bias error in the estimates of ϕ_C , ϕ_{BW} and β. There is, however, some bias error in α, when the true value is very small, arising from the fact that α cannot be negative as it is an RMS quantity.

An example of the bias error and the variability of the four estimates for model 5 is shown in fig. 6.3, which shows the results of applying the 4-parameter model to 32 realizations of the SFOAE output. These are compared against four baseline reference values obtained by applying the 4-parameter model to an ensemble averaged value of the autocorrelation function of the SFOAE data. This shows no significant bias error in the distribution of estimates for any of the four parameters. It also shows that the estimates of ϕ_{BW} are more variable than those of ϕ_C. Note also that the value of α must always be between 0 and 1. The lower limit arises from the fact that it is an RMS quantity. The upper limit arises from the physical constraint of stability. (Note that SOAEs can be present without α exceeding 1, since these only require that $\boldsymbol{rR} = 1$ at a single frequency, rather than that the RMS measure of $\boldsymbol{rR}$ across the frequency range of interest should exceed 1.)

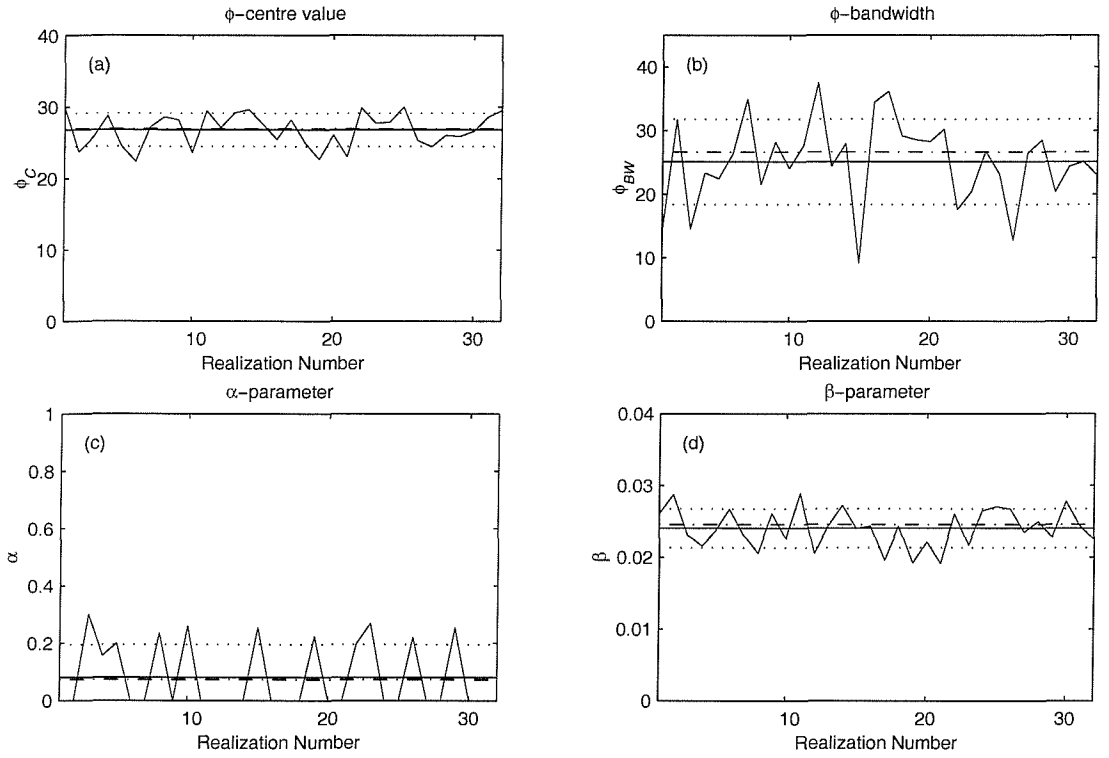


Fig 6.3 Variability in the estimates in the 4 parameter fit to SFOAE data. The 4-parameter model was applied to 32 realizations of the normalised SFOAE, obtained from the cochlear model with a random scattering impedance (model 5). The figures show the mean (thick solid), one standard deviation (thick dahsed) either side of the mean, and the actual estimates for each realization (thin solid). The 4-parameter model was also applied to the ensemble averaged autocorrelation function to obtain ensemble estimates (thick chain) which serve as the baseline reference values. Panel (a): the ϕ-centre value, ϕ_C. Panel (b): the ϕ-bandwidth, ϕ_{BW}. Panel (c): the α-parameter. Panel (d): the β-parameter.

The ability to detect changes in periodicity has also been investigated. This is illustrated in fig 5.4 where changes in periodicity are induced by altering the cochlear amplifier gain. Each panel shows ϕ-spectra at four different amplifier gain settings. Panels (a) and (b) show the induced changes in raw SFOAE ϕ-spectra for two realizations of the SFOAE, whilst panel (c) shows the changes in average SFOAE ϕ-spectrum obtained by ensemble averaging. The corresponding fitted SFOAE ϕ-spectra for these three cases are shown in figs 5.4(e)-(f). Unlike in the raw ϕ-spectra in panel (a), the fitted ϕ-spectra in panel (d), shows a change in peak location with cochlear amplifier gain (albeit a reduced change relative to the ensemble averaged SFOAE ϕ-spectra in panels c and f). This is an example of the increased sensitivity of the estimate of the periodicity of the 4-parameter model, compared to an estimate based on the location of the peak of the raw ϕ-spectrum.

6.9 Summary of the 4-Parameter Model

A 4-parameter model of the normalised SFOAE frequency spectrum has been developed for estimating the ϕ-centre value, and ϕ-bandwidth of the SFOAE ϕ-spectrum. This model is based on the characteristics of a 2^{nd} order Butterworth filter and attempts to account for the spatial filtering proposed by Zweig and Shera (1995) as well as multiple reflections involving the stapes. The performance of the model has been checked against predicted SFOAEs from cochlear models.

7. Predicted Changes in SFOAE Periodicity with the TW Shape

In this section, the 4-parameter model used to quantify the changes in ϕ_C and ϕ_{BW} that cochlear models predict when changes in TW shape are induced.

7.1 The Effect of Cochlear Amplifier Gain on the TW Function

One simple way of altering the TW shape is to progressively deactivate the cochlear amplifier by reducing the gain globally along the entire CP. This has been done using the linear models (model 13-17) as illustrated in fig 7.1 for smooth and active cochlear using either the NK-1986 or the Z-1991 cochlear amplifier. The TW function (magnitude and phase in panels a, b, d and e), has been evaluated against CP place for a single frequency of 1.5 kHz. Panels c and f show the real part of the total CP impedance. Negative values indicate regions of TW amplification. Note that the absence of any scattering impedance is of little importance as the effect of the including a scattering impedance on the TW function is minimal.

Figs 7.1 a, b, d and e show an important result: decreasing the cochlear amplifier gain causes both a broadening of the peak of the TW, and an increase in the TW wavelength in the peak region (which is inversely proportional to the gradient of the phase curve). As discussed earlier, in Shera and Zweig's theory, the wavelength of the TW is related to the periodicity of the SFOAE.

Comparing figs 7.1 a-c with d-f illustrates some obvious differences between the results with the two different cochlear amplifier models. One difference is in the passive response (i.e., where the cochlear amplifier impedance is zero). Model 16, with the T-1998 passive CP, shows a greater accumulated phase at the TW peak (i.e., 5 complete waves) than model 13, with the KdB-1994 passive CP, which shows about 3 complete waves. This is due to the parameter settings rather than any fundamental differences in formulation. A second difference is in the region of negative damping arising from the cochlear amplifier. For the NK-1986 model, the effect of the cochlear amplifier is concentrated in a region just basal to the TW peak. For the Z-1991 model, the cochlear amplifier gives rise to a negative CP resistance spreading apically beyond the characteristic place. It also contributes a significant positive resistance in the more basal regions of the CP.

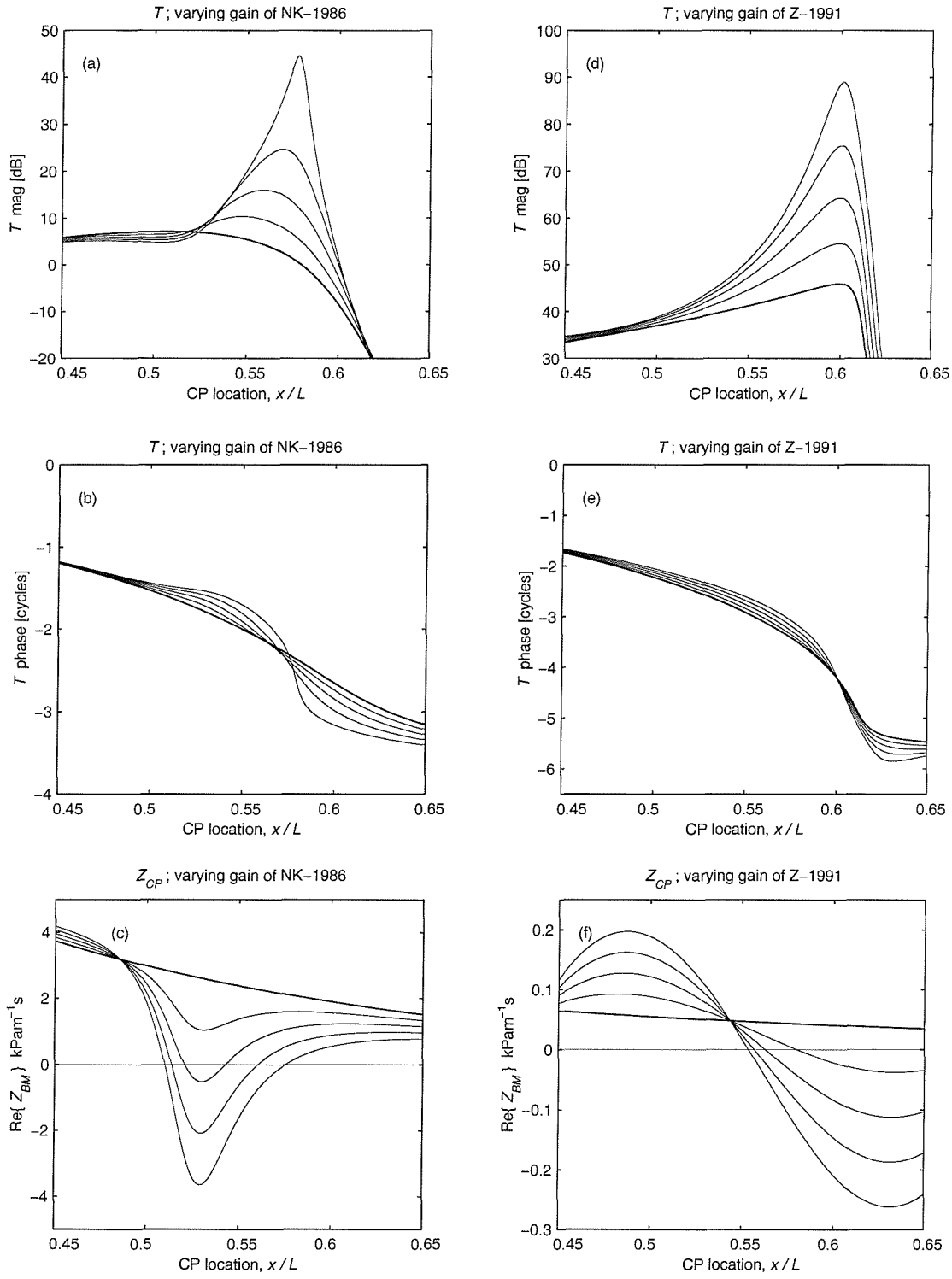


Fig 7.1 The effect of varying cochlear amplifier gain on the travelling wave function, $\boldsymbol{T}$, defined as $v_{CP}(x,\omega)/u_{St}(\omega)$. The travelling wave function is shown against CP location, for a constant stimulus frequency of 1.5 kHz, and is obtained from two active cochlear models (models 13 and 16) with no scattering impedance. Results are shown for cochlear amplifier gain attenuations of 0, 2.5, 6, 12, and ∞ dB. Model 13 has the NK-1986 cochlear amplifier; model 16 has the Z-1991 cochlear amplifier. Panels (a) and (b): magnitude and phase of the travelling wave function for model 13. Panel (c) shows the real part of the total CP impedance (i.e., the resistance) for model 13. Negative resistance indicates TW amplification. Panels (d) - (f) show corresponding results for model 16. The thick line in each panel identifies the result at ∞ dB attenuation (i.e., the passive case).

7.2 The Effect of Cochlear Amplifier Gain on SFOAE Periodicity

The effect on SFOAE periodicity of altering the cochlear amplifier gain has been investigated for three cochlear models, two of which (models 14 and 17) have a random scattering impedance, and a third (model 15), which has a periodic scattering impedance as suggested by Strube. For the random scattering models, results were obtained from an ensemble average over 32 realizations.

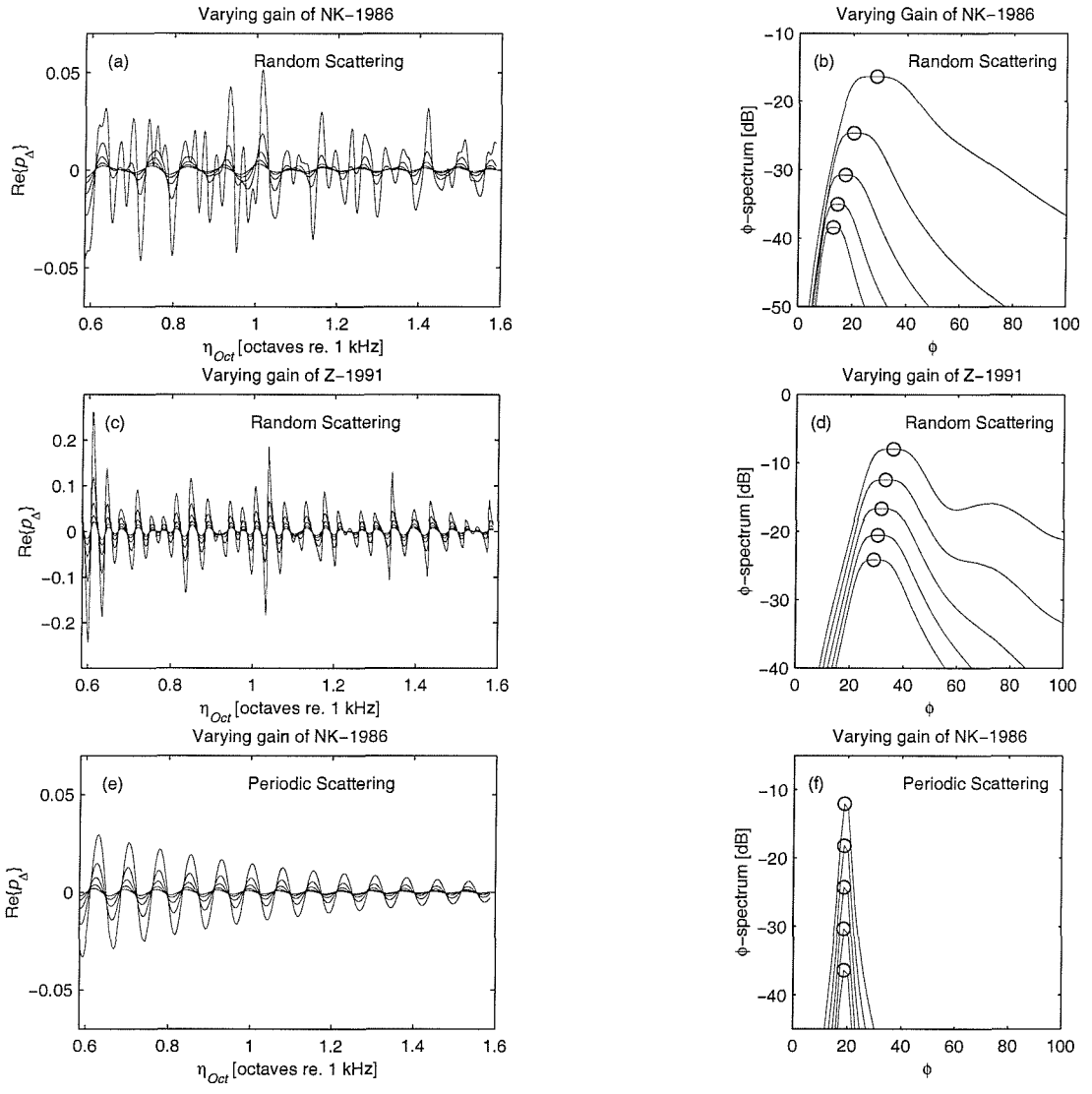


Fig 7.2 The effect of varying cochlear amplifier gain on the normalised SFOAE obtained from three cochlear models (models 14, 17 and 15). Results are presented as single realizations and as fitted ϕ-spectra obtained from the 4-parameter model fit to the ensemble average over 32 realizations. Panels (a) and (b): random scattering impedance with NK-1986 cochlear amplifier attenuated by: 0, 0.4, 0.8, 1.2, 1.6 dB (model 14). Panels (c) and (d): random scattering impedance with Z-1991 cochlear amplifier attenuated by: 0, 0.4, 0.8, 1.1, 1.5, 1.9 dB (model 17). Panels (e) and (f): periodic scattering impedance, with NK-1986 cochlear amplifier attenuated by: 0, 0.4, 0.8, 1.2, 1.6 dB (model 15). Symbol o indicates the peak of the curve.

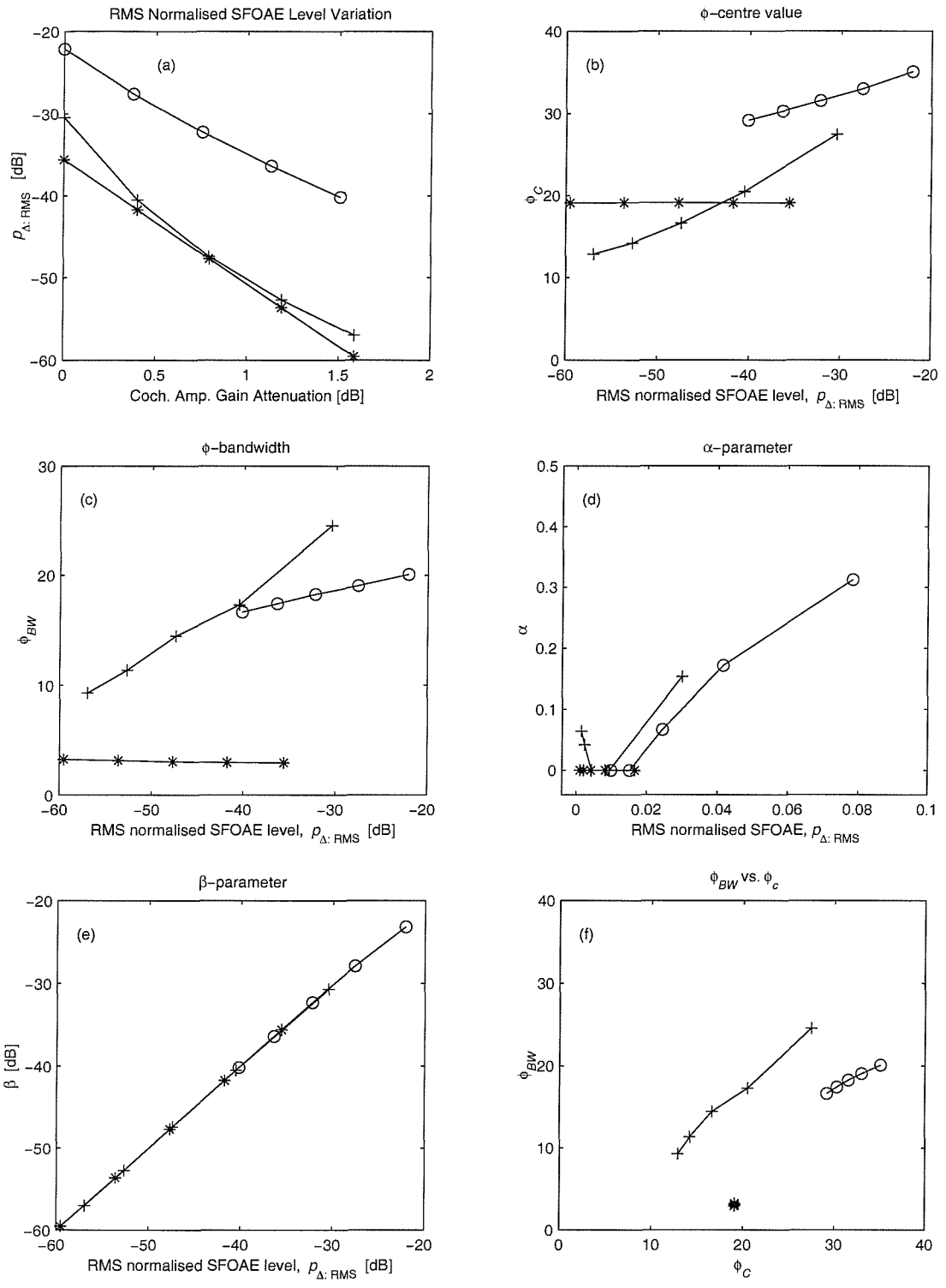


Fig 7.3 The effect of varying cochlear amplifier gain on the SFOAE parameters obtained from three cochlear models (models 14, 17 and 15). Results are obtained by applying the 4-parameter model to the ensemble average over 32 realizations. Model 14 (+ symbol) has random scattering impedance with NK-1986 cochlear amplifier attenuated by: 0, 0.4, 0.8, 1.2, 1.6 dB. Model 17 (o symbol) has random scattering impedance with Z-1991 cochlear amplifier attenuated by: 0, 0.4, 0.8, 1.1, 1.5, 1.9 dB. Model 15 (* symbol) has periodic scattering impedance, with NK-1986 cochlear amplifier attenuated by: 0, 0.4, 0.8, 1.2 and 1.6 dB. Panel (a): the variation of the RMS normalised SFOAE with cochlear amplifier attenuation. Panels (b) - (e): the variation of ϕ_C , ϕ_{BW} , α and β with the RMS normalised SFOAE. Panel (f): the variation of ϕ_{BW} with ϕ_C .

Figs 7.2a – d show the effect of cochlear amplifier gain reduction on single realizations of the SFOAE frequency spectrum for the two random scattering models together with fitted ϕ-spectra obtained by applying the 4-parameter model to the ensemble averaged results. As predicted by Zweig and Shera (1995), for the random scattering models, the ϕ-spectra show an decrease in the ϕ-centre value, ϕ_C, as the cochlear amplifier gain decreases. There is also a corresponding decrease in the ϕ-bandwidth, ϕ_{BW}. In contrast, for the periodic scattering model shown in figs 7.2e and f, no such change in ϕ_C or ϕ_{BW} is seen with changes amplifier gain.

The effect of cochlear amplifier gain on the level of the SFOAE, $L_{\Delta:RMS}$, defined in equation [6.10], is shown in fig 7.3a. Figs 7.3b-e show the four fitted parameters, ϕ_C, ϕ_{BW}, α and β plotted against $p_{\Delta:\,RMS}$. Note that the overall change in the SFOAE level is about 20 dB for each model.

In figs 7.3b and c, the random scattering models show a variation of ϕ_C and ϕ_{BW} with SFOAE level, whilst the periodic scattering model does not. Fig 7.3e shows that β is approximately equal to $p_{\Delta:\,RMS}$, as expected from equation [6.13] when α is small.

In fig 7.3d, α is plotted against the linear normalised SFOAE amplitude, $p_{\Delta:\,RMS}$, (rather than the decibel form) to highlight the approximately linear variation of α with $p_{\Delta:\,RMS}$. This is expected, since, from equation [6.11], $\alpha/\beta = r_0 / g_0$ and $\beta \approx p_{\Delta:\,RMS}$. Therefore $\alpha / p_{\Delta:\,RMS}$ is roughly constant. In these models, r_0 and g_0 are dependent only on the front end models, not on the cochlear amplifier gain. Physically this means that the effect of multiple reflections becomes negligible as the reflectance, $\boldsymbol{R}$, and therefore the SFOAE amplitude, $p_{\Delta:\,RMS}$, becomes small.

Also plotted is the variation of ϕ_{BW} as a function of ϕ_C (fig 7.3f). This gives an indication of the variation of the Q-factor (defined as the ratio of the centre-frequency to the 3 dB bandwidth) of the fitted 2nd order Butterworth filter. For the two random scattering cochlear models, the points lie roughly on a straight line through the origin, indicating a roughly constant Q-factor. The gradient of the slope of this line is the reciprocal of the Q-factor. The differences in Q-factor for the NK-1986 and Z-1991 cochlear amplifiers, where the Q factor is roughly 1.0 and 1.8 respectively, arises from differences both in cochlear amplifier formulation and passive cochlear parameters. For the periodic scattering model, the Q-factor is unrealistically high, at around 6.0, due to the near perfect sinusoidal SFOAE prediction. More realistic values from a Strube model could be obtained by using the Strube narrow band scattering impedance (model 8).

7.3 The Effect of Self-suppression and Two-tone Suppression on SFOAE Periodicity

7.3.1 Self-Suppression Simulations

To model the effect of increasing probe level, the quasilinear models of single tone self-suppression is used (models 18 and 19). This is the model developed by Kanis and de Boer (1993b, 1994, 1996), modified to include reflection sites on the CP. In this model, the cochlear amplifier progressively saturates in some regions of the CP, leading to local reductions in amplifier gain. Figs 7.4 a and b shows the effect of self-suppression due to increasing the stimulus input level on the TW function at a single frequency. Fig 7.4 c shows how (unlike the global gain reduction, fig 7.1c) the saturation of the cochlear amplifier begins at the right hand edge of the active region, where the CP velocity is the highest.

Unlike for the linear models, no ensemble averaging has been performed due to the excessive computation time required. Instead, only two realizations of the random scattering impedance were generated at each probe level. Frequency sweeps were then performed for both the realizations of the random scattering impedance (model 19), and for the corresponding smooth model (model 18). This gave two realizations of the SFOAE frequency spectrum for model 19 at each probe level. The 4-parameter model was then applied to both realizations.

7.3.2 Two-tone Suppression Simulations

As mentioned in section 5.2.3, the predicted effect on the TW shape of introducing a high-side suppressor tone is quite different to that of simply increasing the level of the probe tone (Kanis and de Boer, 1994). This can be seen in fig. 7.4. The reason for this result is that the cochlear amplifier does not saturate uniformly along the CP, but instead tends to saturate first at those points where the amplifier activity is highest. This corresponds to regions where the CP velocity amplitude is highest. Consequently, in self-suppression saturation begins near the peak of the TW of the probe, and then spreads basally as the probe level increases. In contrast, in high-side suppression, saturation can occur near the peak of the suppressor which lies basal to the peak of the probe, and therefore lies within the probe's region of active amplification. Thus the suppressor tone can cause saturation of the cochlear amplifier to begin at a point basal to the peak of the probe's TW envelope, rather than at the peak itself (fig 7.4f).

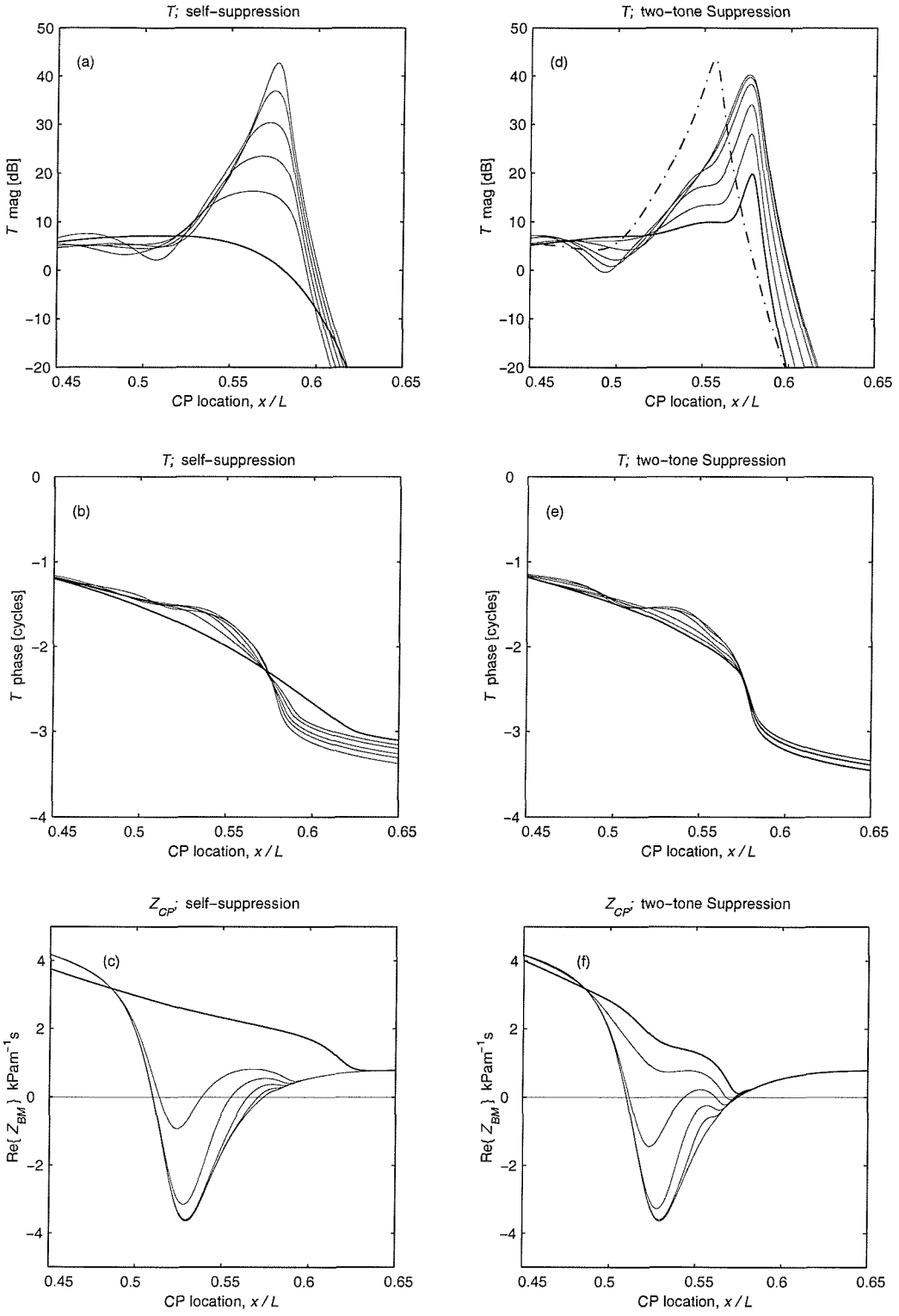


Fig 7.4 The effect of self-suppression and two-tone suppression on the travelling wave function, $\boldsymbol{T}$, defined as $v_{CP}(x,\omega)/u_{St}(\omega)$. The travelling wave function is shown against CP location for a constant stimulus frequency of 1.5 kHz, and is obtained from two quasilinear cochlear models (models 18 and 20) with no scattering impedance. Model 18 simulates self-suppression with probe levels of 40, 50, 60, 70, 80, and 120 dB SPL. Model 20 simulates two-tone suppression with a probe level of 45 dB SPL, and suppressor levels of 30, 50, 60, 70, 80, and 90 dB SPL. Panels (a) and (b): magnitude and phase of the travelling wave function during self-suppression (model 18). Panel (c) shows the real part of the total CP impedance (i.e., the resistance) during self-suppression (model 18). Negative resistance indicates TW amplification. Panels (d) - (f) show corresponding results for two-tone suppression model 20. Thick and thin solid lines: response at stimulus frequency; thick solid lines: response at stimulus frequency in the maximum suppressed case; chain line (shown in (d) only): response at suppressor frequency for 30 dB SPL suppressor level.

Note that the same effect does not occur in low-side suppression, where the peak of the suppressor lies apical to the probe and therefore lies apical to the probe's region of active amplification. Low-side suppression of the probe TW is therefore caused by the CP response to the basal portion of the suppressor TW, rather than its peak response. Consequently, as with self-suppression, saturation of the probe's active amplification begins near the peak of the probe TW, and spreads basally as the low-side suppressor level increases.

The effect of the a high-side suppressor tone on the TW is seen in figs 7.4d-f, for model 20, with a suppressor to probe frequency ratio of 1.1. Compared to the self-suppression case, the changes in TW shape are harder to characterise. It is useful here to define the "6 dB peak region" as the portion of the TW envelope that is no more than 6 dB down from the peak in fig 7.4d. This region shows little if any broadening (fig 7.4d) during two-tone suppression. Similarly the slope of the phase in this region changes relatively little (fig 7.4e). This has consequences for the SFOAE periodicity, since, in Zweig and Shera's theory, periodicity is related to the wavelength of the TW in the peak region, which in turn is related to the slope of the phase of travelling wave. This is changed little by the addition of a high side suppressor tone, whilst it is reduced by a probe level increase.

Note that this definition of the 6 dB TW peak region is somewhat arbitrary. In order to determine precisely how the changes in TW shape seen in fig. 7.4d and e would affect the OAE periodicity directly from Zweig and Shera's phenomenological model (rather than the full cochlear model), the predicted TW response (figs 7.4d and e) could be used to approximate the (scaling symmetric) TW function in equation [4.26]. The periodicity could then be calculated in the manner illustrated in fig. 5.4. This analysis has not been performed. However, provided that the OAE is dominated by reflections from within this 6 dB peak region, the argument stated above that a high-sided two-tone suppressor should cause no reduction in ϕ_C is a valid one.

Two-tone suppression of SFOAEs for various different suppressor levels was simulated using models 20 and 21. The same two realizations of the random scattering impedance that were used in the self-suppression simulation were again used in the two-tone suppression simulation. A frequency sweep was then performed at each suppressor level for both the random scattering impedances (model 21), and for smooth model (model 20), thus yielding two realizations of the SFOAE frequency spectrum. The 4-parameter model was then applied to both realizations.

7.3.3 Results of Self-suppression and Two-tone Suppression Simulations

Figs 7.5a-c show the results of the self-suppression simulations obtained with realization 1. Below a stimulus level of about 40 dB SPL, the model is approximately linear and little self-suppression occurs (panel a). Above this level, the RMS normalised SFOAE level falls off with increasing stimulus level at about 1 dB/dB. Panels b and c show that the changes in the SFOAE ϕ-spectrum are similar to those seen for the global cochlear amplifier gain reduction (fig 7.2b). The reason for this can be seen by examining the spatial variation of the TW phase, the slope of which gives the wavenumber (inversely proportional to wavelength). In both the global gain reduction simulation and in the self-suppression simulation, a reduction in TW peak amplitude is accompanied by an increase in the wavelength near the TW peak, leading to a reduction in the ϕ-centre value, ϕ_C, and hence an to increase in periodicity (equation [4.32]).

Figs 7.5d-f show the results of the two-tone suppression simulations obtained with realization 1. The change in RMS normalised SFOAE level (panel d) shows that two-tone suppression starts to take effect once the level of the suppressor tone exceeds that of the probe tone (45 dB SPL). Unlike either the global gain reduction simulation, a reduction in the SFOAE level is accompanied by a slight increase in the ϕ-centre value, ϕ_C (panels e and f).

The four fitted parameters, ϕ_C, ϕ_{BW}, α and β are plotted against $p_{\Delta:\,\mathrm{RMS}}$ in figs 7.6b-f for realization 1 for both the self-suppression and the two-tone suppression cases. The curve in panel (b), shaped like a lower case "y", clearly shows the difference in the predicted variation of ϕ_C with $p_{\Delta:\,\mathrm{RMS}}$ between the self-suppression and the two-tone suppression cases. The longer limb of the "y" (which has a positive gradient) arises from the self-suppression experiment, and the shorter limb (which has a negative gradient) arises from the two-tone suppression experiment. The difference between the two cases for the variation of ϕ_{BW} is less clear (panel c). Here the curve for the self-suppression case is not monotonic. This is thought to be due to the difficulty in estimating ϕ_{BW}, seen in fig. 6.3b. The explanation for the variation of α and β (panels d and e) is the same as in section 6.2.

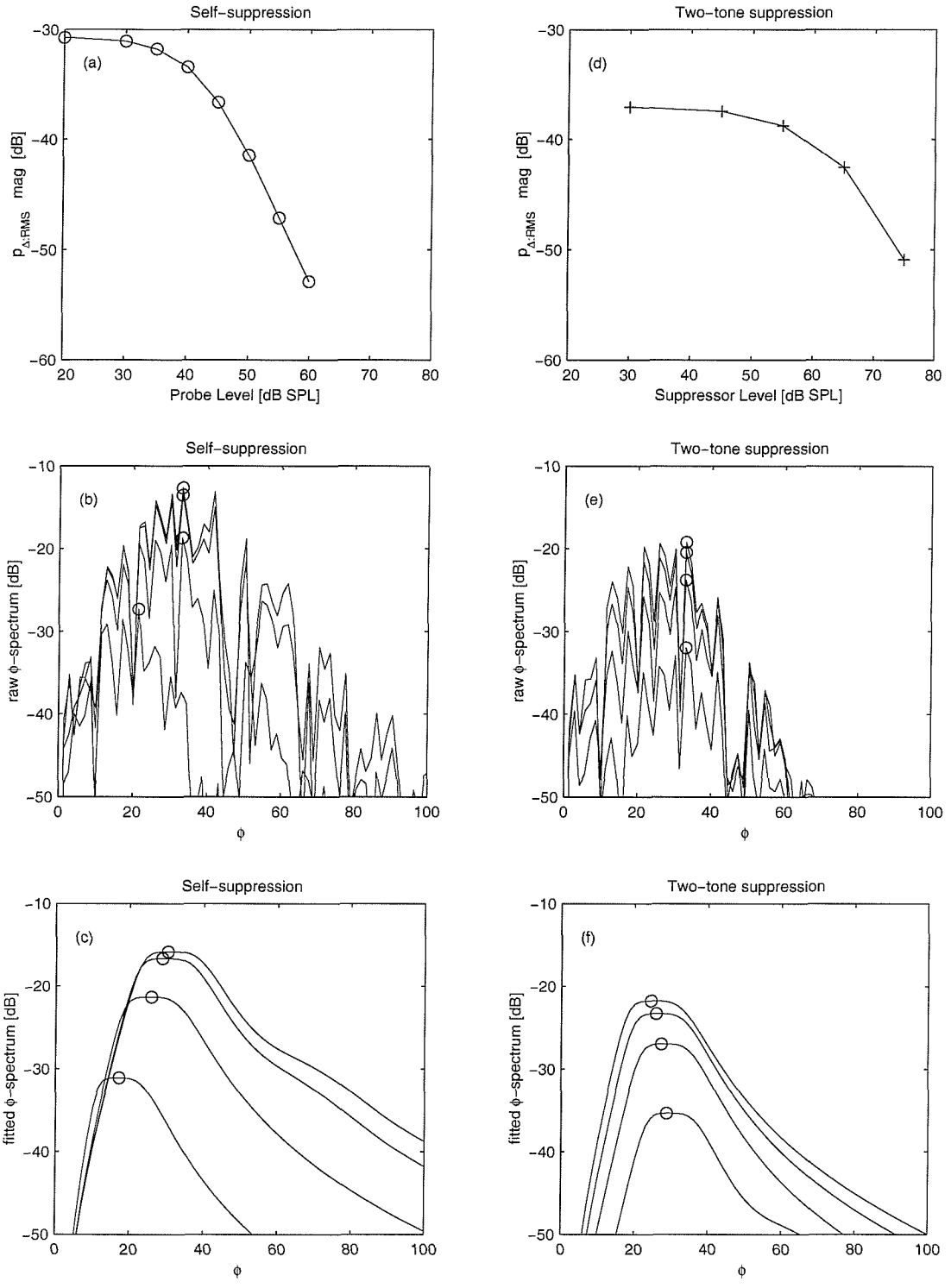


Fig 7.5 The effect of self-suppression and two-tone suppression on the normalised SFOAE and its ϕ-spectrum simulated by quasilinear cochlear models with random scattering impedances (Models 18 and 20). Model 18 simulates self-suppression with probe levels of 20, 30, 35, 40, 45, 50, 55 and 60 dB SPL. Model 20 simulates two-tone suppression with a probe level of 45 dB SPL, and suppressor levels of 30, 45, 55, 65 and 75 dB SPL. Fitted results are obtained by applying the 4-parameter model to a single realization of each frequency sweep. Only single realizations are obtained for each frequency sweep. Panel (a): the variation of the RMS normalised SFOAE with probe level (self-suppression: model 18). Panels (b) and (c): the effect of self-suppression on the raw and fitted SFOAE ϕ-spectra for probe levels of 20, 40, 50, and 60 dB SPL (model 18). Panel (d): the variation of the RMS normalised SFOAE with suppressor level (two-tone suppression: model 20). Panels (e) and (f): the effect of two-tone suppression on the raw and fitted SFOAE ϕ-spectra for suppressor levels of 30, 55, 65, and 75 dB SPL (model 20). In panels (b), (c), (e) and (f) symbol o indicates the peaks of the curve.

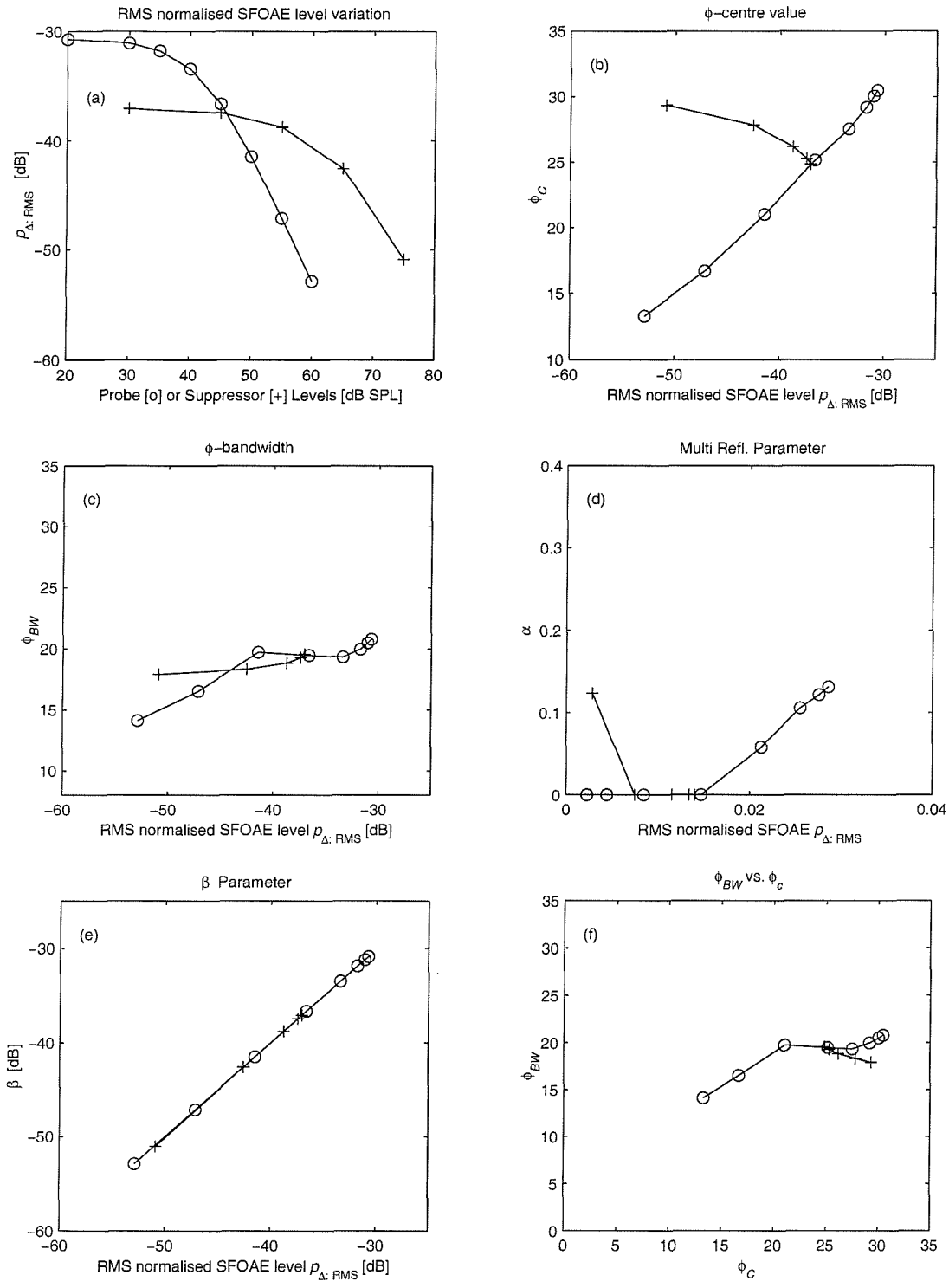


Fig 7.6 The effect of self-suppression and two-tone suppression on the SFOAE parameters obtained from quasilinear cochlear models with random scattering impedances (models 18 and 20). Results are obtained by applying the 4-parameter model to a single realization (realization number 1) for each frequency sweep. Model 18 (o symbol) simulates self-suppression with probe levels of 20, 30, 35, 40, 45, 50, 55 and 60 dB SPL. Model 20 (+ symbol) simulates two-tone suppression with a probe level of 45 dB SPL, and suppressor levels of 30, 45, 55, 65 and 75 dB SPL. Panel (a): the variation of the RMS normalised SFOAE with probe or suppressor level. Panels (b) - (e): the variation of ϕ_C , ϕ_{BW} , α and β with the RMS normalised SFOAE. Panel (e): the variation of ϕ_{BW} with ϕ_C .

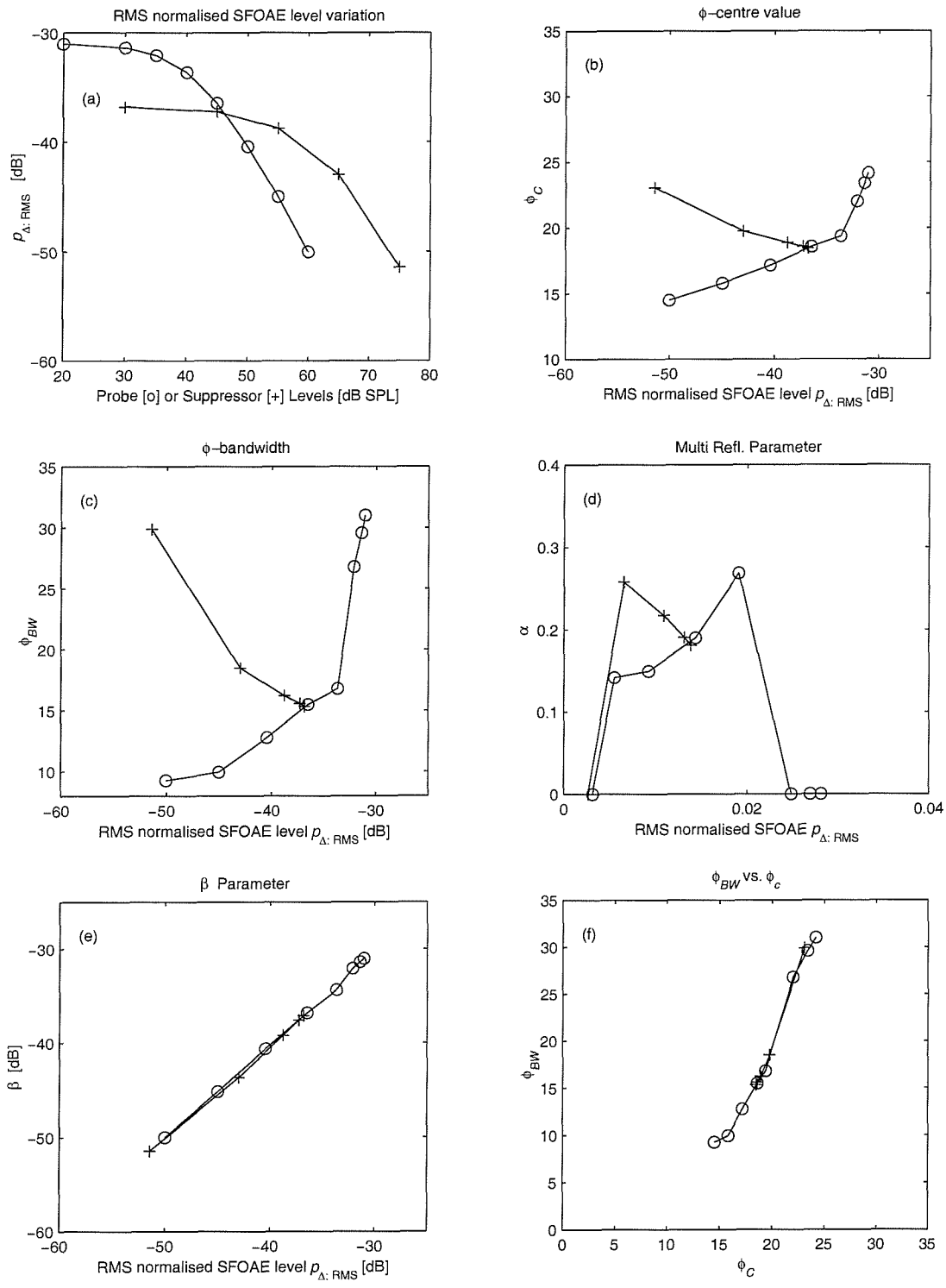


Fig 7.7 As fig. 7.6, except that the simulations used realization 2 of the random scattering impedance instead of realization 1.

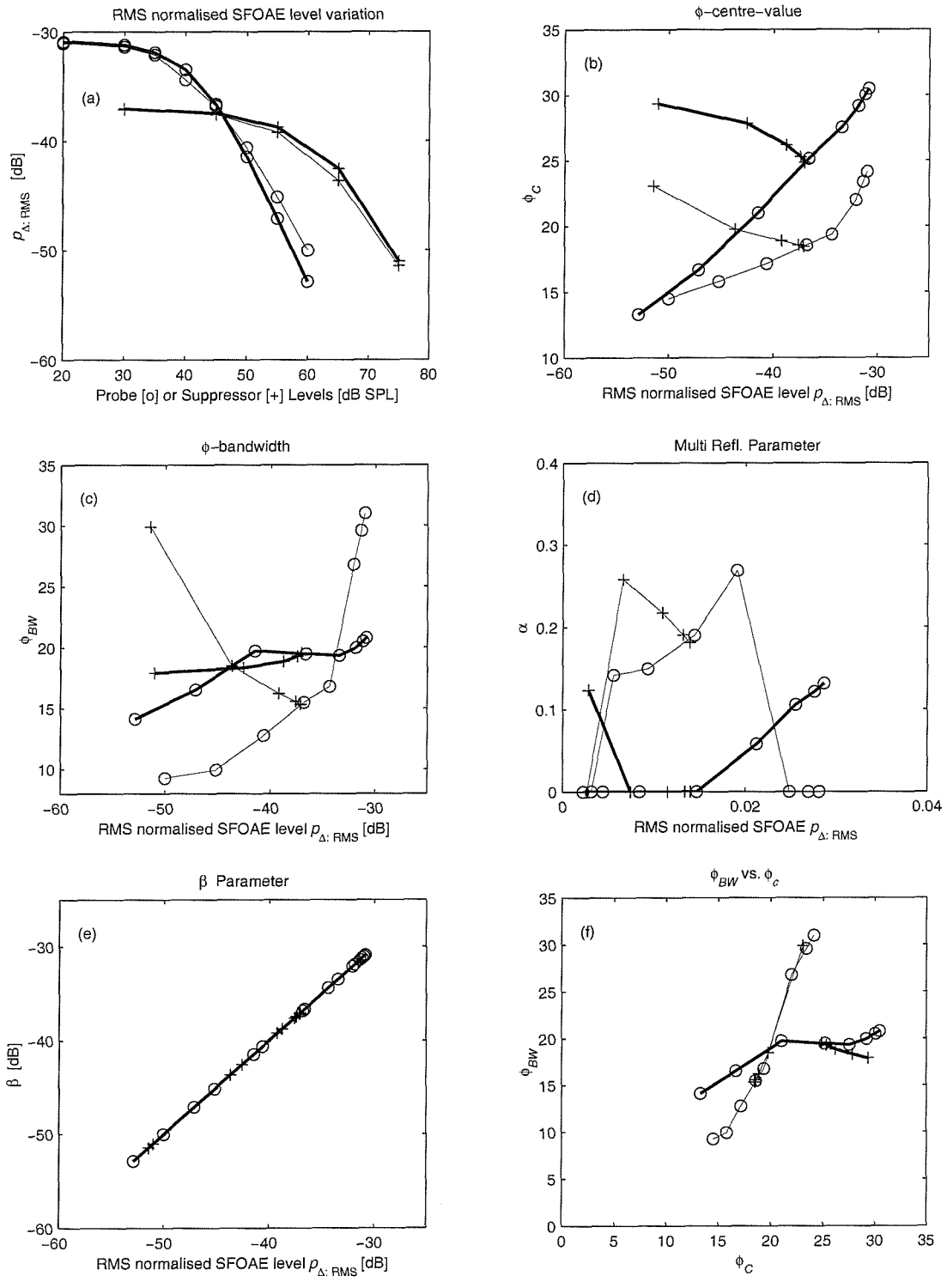


Fig 7.8. Results in figs. 7.6 and 7.7 overlaid. Thick lines: realization 1; thin lines: realization 2.

The corresponding results for realization 2 are shown in fig 7.7a-f. As for realization 1, the variation of ϕ_C with $p_{\Delta:\,RMS}$ follows a y-shaped curve (panel b), though with somewhat lower absolute values of ϕ_C . Unlike realization 1, the variation of ϕ_{BW} now also follows a y-shaped curve. The variation of α (panel d) is somewhat erratic, with several of the points deviating from the expected trend (cf. fig. 7.3d). These results are consistent with the result found in section 6.8, that ϕ_{BW} and α are more difficult to estimate reliably than ϕ_C using the 4-parameter model.

To compare results from realization 1 with those from realization 2, figs 7.8a-f shows both results overlaid. Panel (f) indicates the wide discrepancy in the estimates of the Q-factor of the 2nd order Butterworth filter. In realization 1, the estimated Q-factor is clearly reducing with increasing probe level (i.e., ϕ_C reduces more rapidly than ϕ_{BW}), whilst for realization 2, the reverse is the case. This is a further illustration of the difficulty in estimating ϕ_{BW}.

7.4 Testable Predictions from the Cochlear Models

The results of the cochlear models are qualitatively all in agreement with the predictions of Zweig and Shera's theory of spatial filtering. This theory predicts that, for a cochlea with a random spatial distribution of scattering sites, the form of the SFOAE frequency spectrum is directly related to the shape of the TW. More specifically, the ϕ-centre value, ϕ_C, is related to the wavelength in the TW peak region, whilst the ϕ-bandwidth, ϕ_{BW}, is related to the width of the envelope of the TW peak. Thus the theory predicts that changes in TW shape will result in changes ϕ_C, and ϕ_{BW}.

Changes in TW shape can be induced experimentally by increasing the probe level (self-suppression) or by introducing a suppressor tone (two-tone suppression). Unfortunately changes in the shape of the TW can not be measured experimentally. However, by using the models of self-suppression and two-tone suppression described above, changes in ϕ_C, and ϕ_{BW} can be related to changes in the RMS level of the SFOAE, which give an indication of the degree of suppression. The model results seen in figs 7.6 and 7.7 show clear predictions that may be tested experimentally.

Four testable predictions can be formally stated, based on the characteristic y-shaped results seen in figs. 7.6 and 7.7, panels (b) and (c). These results all apply to the cochlear models based on Shera and Zweig's theory, in which the cochlear scattering impedance has a random spatial distribution, and combined with Kanis and de Boer's quasilinear model of self-suppression and two-tone suppression.

1. For self-suppression, an increase in the level of the probe tone causes a reduction in both the RMS normalised SFOAE and the ϕ-centre value, ϕ_C.

2. For high-sided, two-tone suppression, an increase in the level of the suppressor tone (holding the probe level constant), causes a reduction in the RMS normalised SFOAE, but an increase in the ϕ-centre value, ϕ_C.

3. For self-suppression, an increase in the level of the probe tone causes a reduction in the RMS normalised SFOAE and, in general, a reduction in the ϕ-bandwidth, ϕ_{BW}.

4. For high-sided, two-tone suppression, an increase in the level of the suppressor tone (holding the probe level constant), causes a reduction in the RMS normalised SFOAE. The corresponding value of ϕ_{BW} may increase or decrease. In the latter case, the rate of decrease appears on average to be less than that for ϕ_C.

The results for ϕ_{BW} appear to be more erratic, and show a greater variability between the two realizations than do those for ϕ_C. Thus it might be expected that the predictions for ϕ_{BW} may be less easy to test than those for ϕ_C.

Note that Strube's model, in which the cochlear scattering impedance has a periodic spatial distribution, predicts no change in either the ϕ-centre value or ϕ-bandwidth in either suppression case.

PART III: EXPERIMENTS

8. Experiments to Test the Theoretical Predictions

8.1 Objectives

The objectives of the experiment were to test the four cochlear model predictions stated in section 7.4 against measurements made in human ears. Two experiments, referred to here as *the self-suppression experiment*, and *the two-tone suppression experiment*, have been performed.

The objective of the self-suppression experiment was to determine the variation of the SFOAE amplitude, the ϕ-centre value and the ϕ-bandwidth with the level of the stimulus tone (referred to here as the probe tone).

The objective of the two-tone suppression experiment was to determine the variation of the SFOAE amplitude, the ϕ-centre value and the ϕ-bandwidth with the level of a suppressor tone, at a constant probe tone level.

In both of these, the ϕ-centre value and ϕ-bandwidth were obtained by applying the 4-parameter model to measurements of the SFOAE frequency sweeps.

8.2 Subject Selection

The experiment was designed to compare the results from each subject directly with theory, rather than averaging results across subject. That is to say, from the results of an individual subject, it should be possible to conclude whether or not the theoretical predictions are borne out for that subject. Therefore the number of subjects was chosen to give a representative sample of the normal hearing population (rather than in accordance with any statistical power calculation). It was decided to test one ear per subject, with 20 subjects in total (10 male and 10 female, all aged between 18 and 40 years). Subjects were paid volunteers. The experiment was approved by the departmental Safety and Ethics Committee.

Subjects were screened to ensure that the test ear was normal on otoscopy and tympanometry, and that the pure tone hearing threshold level was 10 dB or better at 0.5, 1, 2, 3, 4 and 8 kHz, and was 15 dB or better at 0.25 kHz.

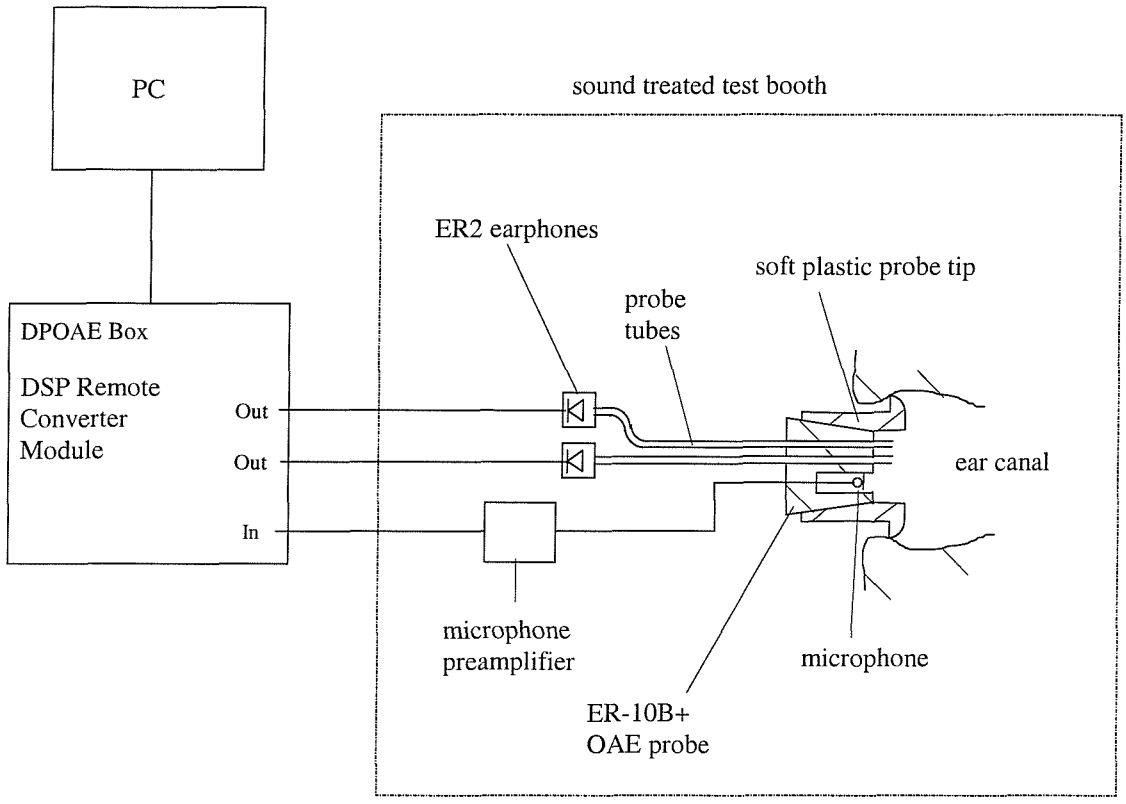


Fig 8.1 Schematic diagram of the experimental apparatus for measuring SFOAEs in the ear canal of a human subject.

8.3 Experimental Apparatus

SFOAEs were measured using an existing system that had been specifically designed for DPOAE and SFOAE measurement, and was controlled by in-house software modified for these experiments. A schematic diagram of the hardware is shown in fig 8.1. The PC controls the DPOAE box, which comprises a DSP card, 16 bit DAC and ADC stages and a sample rate generator. The two outputs of this box connect to two Etymotic ER-2 insert earphones which deliver pure tones to the ear via coupling tubes running through the OAE probe.

Two tones can be presented to the ear simultaneously by the two earphones. One of these tones, termed the *probe* tone, is used to evoke the SFOAE. Thus the SFOAE appears

as a component of the ear canal pressure at the probe tone frequency. The second tone (if present) is termed the *suppressor* tone, and is used to suppress the SFOAE. The term *stimulus* is used to mean the entire acoustic stimulus; that is, the combination of the probe and suppressor tone. Where only one tone is presented, it will still be referred to as the probe tone. (To avoid any confusion, the physical device inserted into the subject's ear canal will always be termed the 'OAE probe', rather the simply the 'probe'.)

The earphones and their tubing provide a high acoustic source impedance, and thus a given earphone voltage gives a volume velocity which is approximately independent of the ear canal impedance loading the OAE probe. The system can deliver tones up to approximately 80 dB SPL in the ear canal. Tones are generated digitally in the DSP card with frequencies which are multiples of 16 Hz. The acoustic pressure in the ear canal is measured using a low noise Etymotic ER-10B+ microphone housed in the OAE probe, and connected through a preamplifier to the input of the DPOAE box. The OAE probe is connected to the ear canal using a soft plastic cuff.

The OAE probe is calibrated in a Bruel and Kjaer Type 4157 ear simulator which contains a reference microphone situated at a point corresponding approximately to the eardrum. Earphone and microphone levels quoted in all the experiments are based on this calibration. Thus a nominal earphone level of 20 dB SPL is defined as the voltage amplitude to the earphone that gives 20 dB SPL at the reference microphone of the ear simulator. Similarly, an OAE probe microphone level of 20 dB SPL is defined as the voltage output from the OAE probe microphone that arises when the reference microphone of the ear simulator is reading 20 dB SPL, when excited by the OAE probe earphones. Calibration was performed at 16 Hz intervals over the frequency range of interest. It can be shown that the problem of standing waves in the ear simulator, which is inherent to this method for defining microphone calibration at high frequencies (>3 kHz), does not materially influence the results of the SFOAE experiments conducted here.

The experiment was conducted with the subject sitting in a sound treated booth, whilst the researcher, the DPOAE box and the PC are located outside the booth.

8.4 Measurement of Ear Canal Pressure during Stimulus Presentation

At any probe frequency, the SFOAE is calculated from two measurements of the ear canal pressure: the first in which both the SFOAE and the stimulus components are present, and the second in which the SFOAE component has been minimised by acoustic suppression. The details of this method are given in sections 8.8.1 and 8.8.2.

The method of measurement of the ear canal pressure at a given probe frequency provides for both noise reduction by synchronous averaging, and for the rejection of noisy epochs of the signal. The stimuli are delivered in epochs of 62.5 ms (2048 sample points) containing a whole number of periods of the stimulus tones. In the DSP card, following the ADC[32] stage, the sampled microphone signal is recorded in corresponding epochs with 16 bit precision and loaded into buffers of 2048 points. On completion of the epoch and before the next epoch begins, an in-place FFT with 16 Hz resolution is calculated. The buffer holds a whole number of periods of the tones, thereby eliminating any truncation effects in the FFT. The FFT is transferred to the PC for further processing as follows.

The first buffer (i.e., the first 62.5 ms of the signal at each frequency) is always discarded, since this will contain the transient response of the transducers and the ear. Subsequent buffers are then classified as 'good' or 'bad' depending on an estimate of their noise level, as described below. Bad buffers are rejected, whilst good buffers are accepted for averaging.

Averaging of the (complex) FFT across successive buffers leads to an improvement in SNR equal to 3 dB for each doubling of the number of buffers in the average (assuming the noise is gaussian and uncorrelated with the stimulus). The number of good buffers in the average varied between 8 and 128, depending on estimates of the levels of SFOAE signal and noise. Thus the total duration of the 'good' section of signal varied between 0.5 and 8 seconds.

The buffer was classed as 'bad' if the estimated noise power in the buffer around the frequency of interest (i.e., the probe frequency) exceeded some preset threshold, referred to here as the rejection level. This allows particularly noisy sections of data (perhaps arising because the subject has moved or swallowed) to be rejected before they could contaminate the average. The noise level is estimated by averaging the power in the microphone signal in the 10 spectral lines either side of the probe spectral frequency line (but excluding the spectral line at the location of the suppressor tone). The assumption here was that the noise power in this 16-Hz band equalled the average noise power in the 20 adjacent spectral bands. The setting of the rejection level for the two experiments is discussed further in section 8.8.3. Typically the rejection level was set to 10 or 15 dB SPL. Note also that the noise is subject to truncation effects, leading to spurious components in the FFT.

[32] The ADC is an 8 × over-sampling device, thus eliminating the need for anti-aliasing filters.

8.5 Measurement of the SOAEs

In order to identify any SOAEs, the power spectrum of the ear canal pressure in the absence of any stimulus was measured. Since SOAEs are not phase locked to the measuring epoch, synchronous averaging is not appropriate. Instead, the average power in each 16-Hz band was calculated over 100 buffers. As in the previous section, 'bad' buffers are rejected. Typically the rejection level was set at about 15 dB SPL. The frequency for calculating the noise level used in the rejection procedure was arbitrarily chosen as 1 kHz.

This method of measuring SOAEs is relatively insensitive, because the frequency resolution of the power spectrum (16 Hz) is poorer than the bandwidth of a typical SOAE, which may be as narrow as 1 Hz or less. The consequences of this are discussed in a later section.

8.6 Experimental Procedure

Prior to testing, subjects were screened for normal hearing. Suitable subjects were then seated in the test booth, and the OAE probe inserted into the test ear. Subjects were instructed to keep as still as possible, and to swallow as infrequently as was comfortable whilst testing was in progress. The test session was then carried out.

For each test session, two experimental protocols were followed: the self-suppression and the two-tone suppression protocol. Both protocols comprised initial measurements of SOAEs, followed by a series of frequency sweeps, followed by a final measurement of SOAEs. The series of frequency sweeps will be referred to as a 'level series'. In the self-suppression protocol, described in detail in section 8.8.1, the level series comprised seven frequency sweeps, each at a different probe level. In the two-tone suppression protocol (section 8.8.2), the level series comprised five frequency sweeps, each at a the same probe level, but at different suppressor levels. The first of these frequency sweeps was performed with the suppressor absent. Each frequency sweep comprised measurements of the SFOAE pressure (referred to as 'points') at 91 frequencies between 1376 Hz and 2816 Hz. Each SFOAE pressure measurement is obtained from two measurements of the (complex) ear canal pressure at the probe frequency. The first of these two measurements yields the stimulus + SFOAE pressure component, whilst the second yields an estimate of the stimulus component alone. This method is described in detail in the next section. As a check on repeatability, each subject sat for two test sessions on different days.

The self-suppression and two-tone suppression protocols took roughly 50 and 30 minutes respectively to carry out. Removal and refitting of the OAE probe was kept to a minimum. Usually this meant that the OAE probe remained in place for the entire duration of the session. As a minimum, the OAE probe was kept in place for the duration of each frequency sweep (i.e., sweeps were discarded if the OAE probe was removed before completion).

For the first session, the SFOAE frequency sweeps were collected in order of increasing probe (or suppressor) level. This order was reversed for the second test session (though, as expected, the order of presentation made no noticeable difference).

Summarising the basic test structure, and specifying the main experimental parameters:

- Each subject sat for 2 sessions.
- For each session, two experimental protocols were followed: the self-suppression and two tone suppression protocol.
- Each protocol comprised initial SOAE measurements, followed by one SFOAE level series, followed by final SOAE measurements.
- The self-suppression level series comprised seven frequency sweeps at probe levels ranging from 14 up to 50 dB SPL in 6-dB steps.
- The two-tone suppression level series comprised five frequency sweeps: one unsuppressed sweep plus four suppressed frequency sweeps at suppressor levels ranging from 26 up to 62 dB SPL in 12-dB steps. The probe level was held constant at 26 dB SPL. The suppressor frequency was $1.3 \times$ the probe frequency.
- Each frequency sweep comprised 91 points at frequencies from 1376 to 2816 Hz in 16-Hz steps.
- Each point comprised a single measurement of the SFOAE pressure obtained from two measurements of the complex ear canal pressure at the probe frequency: the first is the stimulus + SFOAE components, the second is an estimate of the stimulus component alone.
- Each measurement of the complex ear canal pressure was obtained by averaging a number of the FFTs (between 8 and 128), each FFT being calculated from a recording epoch with a duration of 62.5 ms.

8.7 The Problem of the Frequency Resolution and Spectral Aliasing

One question that must be addressed is whether or not the frequency resolution of 16 Hz (which was set by the DSP card) was sufficient to accurately characterise the SFOAE frequency spectrum. This can be thought of as a 'spectral aliasing' problem, analogous to the more familiar problem of temporal aliasing. Temporal aliasing arises during analogue to digital conversion when the analogue signal contains frequency components that are higher than half the sampling rate (the so called 'folding frequency'). The problem is not just that these high frequency components cannot be represented by the digital time series, and are therefore lost, it is that they appear as spurious low-frequency components. Once sampled, they are indistinguishable from genuine low-frequency components, and no amount of digital signal manipulation can remove them (e.g., Newland, 1987).

The procedure of measuring the SFOAE frequency spectrum at discrete intervals is, in effect, sampling the true underlying continuous spectrum, and therefore susceptible to the problem of aliasing. Thus, any variations in the continuous frequency spectrum of the SFOAE that occur over sufficiently small frequency intervals will be misinterpreted in the measured discrete spectrum as variations over larger frequency intervals. To assess this problem, it is first assumed that the SFOAE is roughly stationary in the η-domain (section 6.5). Then to use the analogy of temporal aliasing, consider η as the analogue of time, and ϕ as the analogue of frequency. Aliasing becomes a problem when components in the true ϕ-spectrum are higher than the 'folding' value of ϕ. The analogy is not perfect, because the measurements are made at constant frequency intervals of 16 Hz, rather than constant η-intervals. The intervals in η_{Oct} decrease over the measurement range of 1376 to 2816 Hz. However, if we assume (pessimistically) that the η-series is obtained with the largest η_{Oct} interval then we obtain a sampling interval of:

$$\Delta\eta_{Oct} \equiv \log_2[(1376+16)/1376] = 0.0167 \text{ octaves}$$

and therefore a 'folding' ϕ value of:

$$\phi_{\text{fold}} = 1/\Delta\eta = 1/(\log_e 2\,\Delta\eta_{Oct}) = 1.443 / 0.0167 = 86.$$

This means that components with ripple spacings down to a periodicity value of 1/86 can be represented, and will not cause aliasing (recall that typical periodicities are 1/15). By examining the measured raw ϕ-spectra of the frequency sweeps (to be presented later), it is clear that the spectrum at $\phi = 86$ is at least 20 dB below the maximum value, and therefore spectral aliasing will not be a problem.

Two further points are worth noting. The first is that here the 'folding' ϕ is equal to the 'sampling' ϕ, rather than half the 'sampling' ϕ, where the latter value might be expected from the temporal aliasing analogy. The reason for this is that the SFOAE η-series is complex and has a ϕ-spectrum that is entirely one sided. Therefore the ϕ-spectrum has no left hand side to 'fold back' onto the right hand side (Newland, 1987). The second point is that, with a constant frequency interval, spectral aliasing becomes more of a problem at lower frequencies. In fact, the minimum frequency value of 1376 Hz used in this experiment was chosen partly to avoid possible spectral aliasing problems.

8.8 SFOAE Measurements

8.8.1 The Self-suppression Experimental Protocol

The measured SFOAE is a function of two independent variables: the probe input pressure amplitude, and probe input frequency. For the self-suppression experiment, SFOAEs frequency functions were measured at each of seven different input levels over a frequency range of just over one octave. The seven input levels used were from 14 up to 50 dB SPL in 6-dB steps. For each of the seven SFOAE frequency sweeps, the probe input level was held constant while the frequency was decreased from 2816 to 1376 Hz in 16-Hz steps. Note that the probe input pressure amplitude is not controlled directly. Instead, the voltage to the earphone is controlled based on the level in the B & K Type 4157 ear simulator (section 8.4). Thus, for example, an input level of 20 dB SPL means that the voltage to the earphone (essentially proportional to the earphone volume velocity) is that which gives 20 dB SPL at the reference microphone in the ear simulator.

Calculating the SFOAE at any frequency and input level requires using the OAE probe microphone to measure the complex ear canal pressure in two conditions. In the first condition, ear canal pressure equals the stimulus pressure component plus the OAE pressure component. In the second it is approximately equal to the stimulus pressure component alone. The SFOAE is then the (complex) difference between these two. To estimate the stimulus component alone, the suppressor method of SFOAE measurement was used, where a high-level suppressor tone is introduced to reduce the OAE component as far as possible. A suppressor tone at 58 dB SPL, and spaced 16 Hz above the probe tone was found to provide adequate suppression. For example, to obtain the SFOAE component for an input level of 20 dB SPL, at 2000 Hz, the following pair of measurements is made:

Table 8.1 Example of a pair of presentations for a single point in the SFOAE frequency sweep for the self-suppression experimental protocol

Mic. Components	Probe Tone		Suppressor Tone	
	freq (Hz)	level (dB SPL)	freq (Hz)	level (dB SPL)
$p_{EC} = p_{EC:R=0} + p_{SF}$	2000	20	==== None ====	
$p_{EC:R=0}$	2000	20	2016	58

The two measurements must be completed within a short time interval (e.g., two seconds or so), in order to minimise 'drift noise' discussed further in section 8.8.3. The measurement of the ear canal pressure is detailed in section 8.4.

8.8.2 The Two-tone Suppression Experimental Protocol

Unlike the self-suppression experiment, in the two-tone suppression experiment, SFOAEs are measured in the presence of a suppressor tone. Such SFOAEs will be called partially suppressed SFOAEs (to distinguish them from the maximally suppressed SFOAEs used to estimate the stimulus component alone).

The partially suppressed SFOAE is a function of four independent variables. These are the input pressure amplitude and frequency of both the probe and suppressor tones. For the two-tone suppression experiment, four SFOAE frequency sweeps were measured, each with the same probe input level of 26 dB SPL, but with four different suppressor levels (from 26 to 62 dB SPL in 12-dB steps). The suppressor frequency was nominally 1.3 times the probe frequency (rounded to the nearest 16 Hz), whilst the probe frequency was decreased from 2816 to 1376 Hz in 16-Hz steps. As before, each of the four SFOAE frequency sweeps was measured by varying the frequency while the probe and suppressor tone levels were held constant. In addition to the four partially suppressed SFOAE frequency sweeps, a fifth, unsuppressed measurement was made, also with a 26 dB SPL probe. This was an exact repeat of one of the self-suppression frequency sweeps, but was included to give a baseline measurement made close in time to the four partially suppressed sweeps.

The partially suppressed SFOAE is measured using a similar method to that for unsuppressed SFOAEs, except that a suppressor tone is introduced. For example, to obtain the (partially suppressed) SFOAE component for a probe tone of 26 dB SPL and 2000 Hz,

in the presence of a suppressor tone of 38 dB SPL and 2608 Hz, the following pair of measurements are made:

Table 8.2 Example of a pair of presentations for a single point in the SFOAE frequency sweep for the two-tone suppression experimental protocol

Components	Probe Tone		Suppressor Tone	
	freq (Hz)	level (dB SPL)	freq (Hz)	level (dB SPL)
$p_{EC} = p_{EC:R=0} + p_{SF}$	2000	26	2608	38
$p_{EC:R=0}$	2000	26	2016	58

Note that the purpose of the suppressor is different in the two measurements: in the first its purpose is to partially suppress the SFOAE by a controlled amount, whilst in the second, it is to suppress the SFOAE as far a possible.

8.8.3 Optimising the Signal-to-noise Ratio in SFOAE measurements

The measured SFOAEs are small relative to the background noise, and therefore great care was taken to achieve an adequate signal-to-noise ratio (SNR) in the design of the experiment. In the following discussion the term 'noise' is used to refer to the uncertainty in the measurement of the SFOAE, rather than to any actual physical contaminating signal. Specifically, the noise on the SFOAE signal is defined here as the standard deviation in the estimate of the SFOAE pressure, and is denoted by n_{SF}. With this definition, two different sources of noise can be identified. These will be called here 'additive noise' and 'drift noise', and are denoted by $n_{SF:Add}$ and $n_{SF:Drift}$ respectively. These components are all functions of probe frequency, f_1. The SNR for the SFOAE is then defined as:

$$\mathrm{SNR}(f_1) \equiv 10\log_{10}\frac{|p_{SF}(f_1)|^2}{n_{SF}^2(f_1)} \qquad [8.1]$$

Where in [8.1], p_{SF} is understood to refer to the true value of the SFOAE, rather than the actual measurement, which will include noise.

Additive noise arises from unknown random components in the microphone signal, which may originate from acoustic sources other than the cochlea (such as physiological noise) or from electrical noise in the electronic components. This additive noise is reduced by the signal averaging procedure, such that the noise power reduces by approaching 3 dB for each doubling of the number of buffers in the average. It can easily be shown that,

since the error in the SFOAE arises from the noise in two measurements of the ear canal pressure, the additive noise component is given by:

$$n^2_{SF:Add}(f_1) = n^2_{EC}(f_1) + n^2_{EC:R=0}(f_1) \qquad [8.2]$$

where n_{EC} and $n_{EC:R=0}$ are the standard deviations of the measurements of the ear canal pressure in the unsuppressed and suppressed conditions, respectively. Good estimates of these two values can be obtained from the method of averaging the noise power in adjacent spectral bands. The values of n_{EC} and $n_{EC:R=0}$ are usually approximately equal and therefore the power of the SFOAE additive noise is twice that of the additive noise in a single ear canal pressure measurement.

Drift noise arises from the calculation of the SFOAE pressure as the difference in two microphone readings: $p_{SF} = p_{EC} - p_{EC:R=0}$ (first introduced in equation [4.30]), and from the fact that these two readings cannot be obtained simultaneously. Slow variations in these components appear as spurious SFOAE readings. Now the estimate of the SFOAE is calculated from a small difference in two large quantities, and is therefore highly sensitive to small changes in the measured ear canal pressure. The effect can be illustrated by the following example. Consider measurements in an ear with no SFOAEs. Two measurements, about 1 second apart, are made of the ear canal pressure, the first in the unsuppressed condition and the second in the suppressed condition. Any change in the two measurements is attributed to the presence of an SFOAE in the unsuppressed condition. However, changes will also arise from other sources such as a change in the middle-ear pressure, or a change in the OAE probe fit over the course of the two readings. In addition low frequency additive noise (with a period greater than 1 s) will also lead to changes between the two readings. Consider for example, a change in the source impedance due to a change in probe fit. This will cause a change in ear canal pressure measured at the microphone for a given earphone voltage. Furthermore, for a given change in impedance, the size of the change in ear canal pressure will be proportional to the original ear canal pressure. This means that the magnitude of the drift error increases as the stimulus level increases. Because the SFOAE pressure is relatively smaller at high levels, the drift noise becomes more significant as the level increases. For example, at the highest stimulus levels, the SFOAE level is typically 40 dB below the stimulus level. Therefore a drift of 1% (or 0.09 dB) in the amplitude of the measured ear canal pressure leads to an error in the SFOAE pressure amplitude of 100% (or 6 dB).

Drift noise is also increased as the time interval between the two readings is increased. Therefore, unlike additive noise, the drift noise is increased, rather than reduced, by increasing the averaging time for the signal: the longer the time interval, the

greater the opportunity to drift. Drift error could be reduced by an ensemble averaging technique in which the two ear canal measurements are made repeatedly at each frequency. However this technique was found to be too time consuming to pursue.

In measuring the SFOAE, a trade-off exists between additive noise and drift noise. The additive noise is caused by noise on the microphone signal at the probe frequency, and can be reduced by increasing the number of averages. In magnitude, it is independent of the stimuli, and it is therefore most serious at low probe levels where the SFOAE pressure is small. The drift noise arises from slow fluctuations in the microphone signal and can be reduced by reducing the time interval between the measurement of the unsuppressed and suppressed ear canal pressure components. It is most serious when the SFOAE pressure is small relative to the probe pressure, as occurs at high probe levels. Thus increasing the averaging time reduces the additive noise, but increases the drift noise. The two main recording parameters: rejection level and averaging time were set such that, on average, the total SNR for the SFOAE was roughly independent of the stimulus conditions. This was achieved by averaging for longer time periods at low probe levels, where the additive noise is dominant, than at high probe levels, where drift noise is dominant. An exception to this was at the very lowest probe levels, where additive noise is dominant, but where time limitations made it impossible to achieve the desired SNR. This is because each improvement in SNR of 3 dB requires a doubling of averaging time, leading to excessive recording times.

9. Results

The results were processed to obtain the following for each subject:

1. The variation of the fitted SFOAE ϕ-spectrum (in particular of the ϕ-centre value, ϕ_C, and ϕ-bandwidth, ϕ_{BW}), and the RMS normalised SFOAE level during self-suppression and two-tone suppression.
2. Estimates of the SNR.
3. A measure of the influence of SOAEs on the results.

The procedures for processing these results are discussed below.

9.1 Calculating the Fitted SFOAE ϕ-spectrum from Measured Data

For each subject, there are data for two test sessions; for each session, there are frequency sweep data for seven different probe levels, and for five two-tone suppressor levels. For each of these frequency sweeps the following steps were carried out:

1. The SFOAE frequency spectrum, $p_\Delta(f)$ was obtained from the raw data.
2. The function was re-sampled from equal frequency to equal log-frequency intervals, to obtain the η-series, $p_\Delta(\eta)$.
3. The estimate of the autocorrelation function, $\tilde{R}_{p_\Delta p_\Delta}(\eta')$, was calculated.
4. The 4-parameter model was applied to give the estimates of ϕ_C, ϕ_{BW}, α and β which define the fitted SFOAE ϕ-spectrum.

These steps are detailed below. Fig. 9.1 illustrates the steps for an example frequency sweep, taken from the measurements in subject no. 20, session 1. These results are typical for subjects with moderate to strong SFOAEs.

Step 1

As explained in section 8.8.1 and 8.8.2, for each probe frequency, f_1 , the ear canal pressure was measured by the OAE microphone in two conditions. In the first condition, the ear canal pressure is denoted $p_{EC}(f_1)$ and is measured either with no suppressor tone (self-suppression experiment) or with a suppressor tone present (two-tone suppressor experiment). In the second condition, the ear canal pressure is denoted $p_{EC:R=0}(f_1)$, and is measured in the presence of the maximum suppressor tone. This gives an estimate of the stimulus pressure component (i.e., the ear canal pressure in the non-reflecting cochlea condition). Results for the self-suppression and two-tone suppression experiments were processed in the same way.

In step 1, an estimate of the SFOAE pressure, $p_{SF}(f_1)$, is obtained. This quantity will be termed the 'measured SFOAE pressure' (though it should be borne in mind that this is not measured directly, but is only an estimate of the true SFOAE, based on the two measurements of ear canal pressure). This is normalised with the ear canal pressure measured in the maximally suppressed condition, $p_{EC:R=0}(f_1)$, to give the normalised SFOAE pressure, $p_\Delta(f_1)$, (denoted by Δ in Zweig and Shera, 1995).

$$p_{SF} \equiv p_{EC} - p_{EC:R\approx 0} \qquad [9.1]$$

$$p_\Delta \equiv \frac{p_{SF}}{p_{EC:R\approx 0}}$$

where all quantities are complex and functions of the probe frequency, f_1 (91 values at 16-Hz intervals).

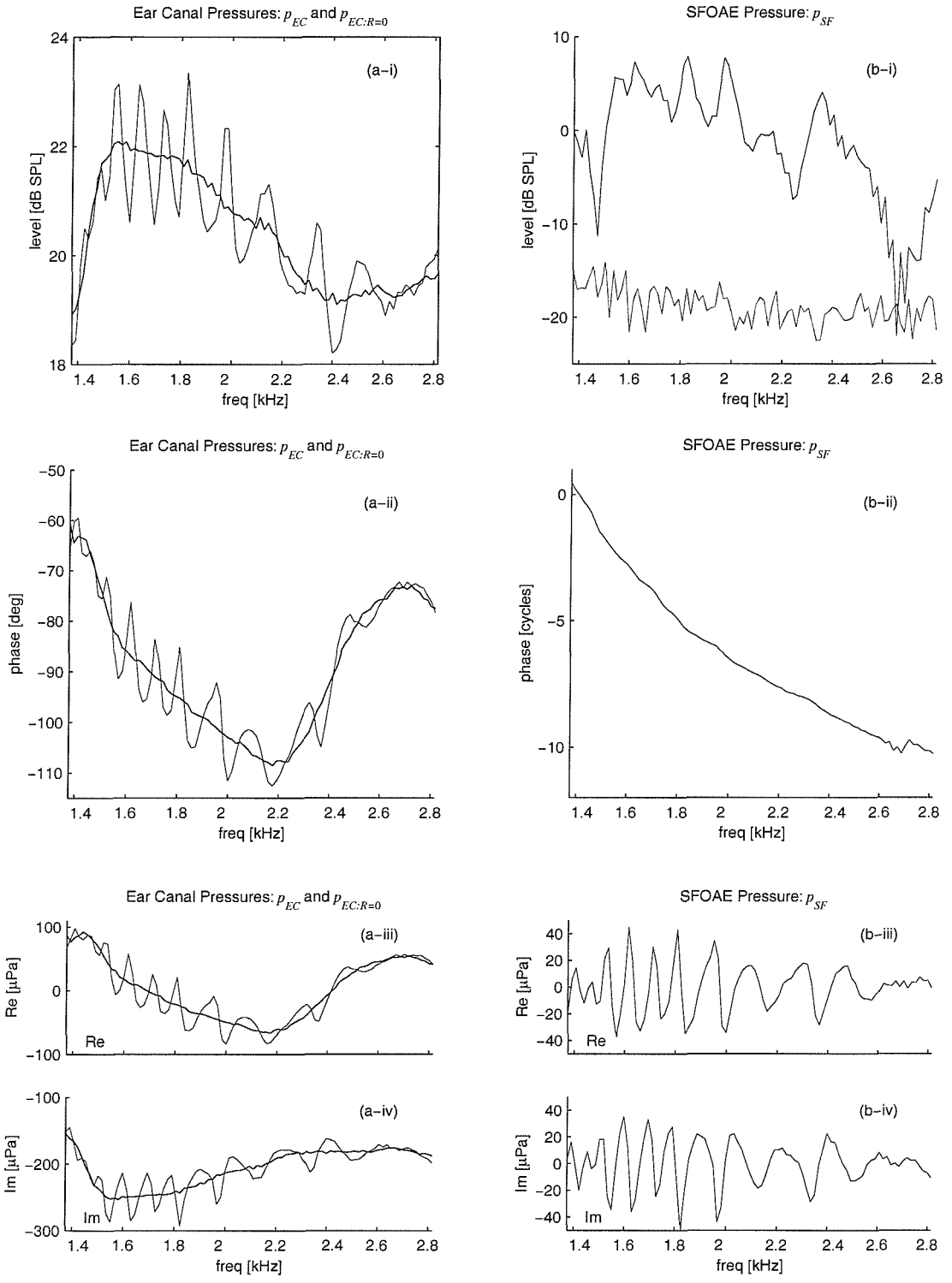


Fig 9.1 Signals at each stage in the processing of measured ear canal data to give the fitted SFOAE ϕ-spectrum. (Data from subject 20, session 1, earphone level 20 dB SPL nominal, no suppressor). Panels (a-i) -(a-iv): magnitude, phase, real and imaginary parts respectively, of both the unsuppressed ear canal pressure, $\boldsymbol{p_{EC}}$, (thin line) and the maximally suppressed ear canal pressure, $\boldsymbol{p_{EC:\,R=0}}$ (thick line). Panels (b-i) -(b-iv): magnitude, phase, real and imaginary parts respectively, of the SFOAE pressure, $\boldsymbol{p_{SF} \equiv p_{EC} - p_{EC:\,R=0}}$. The thin line in (b-i) is the estimated noise floor.

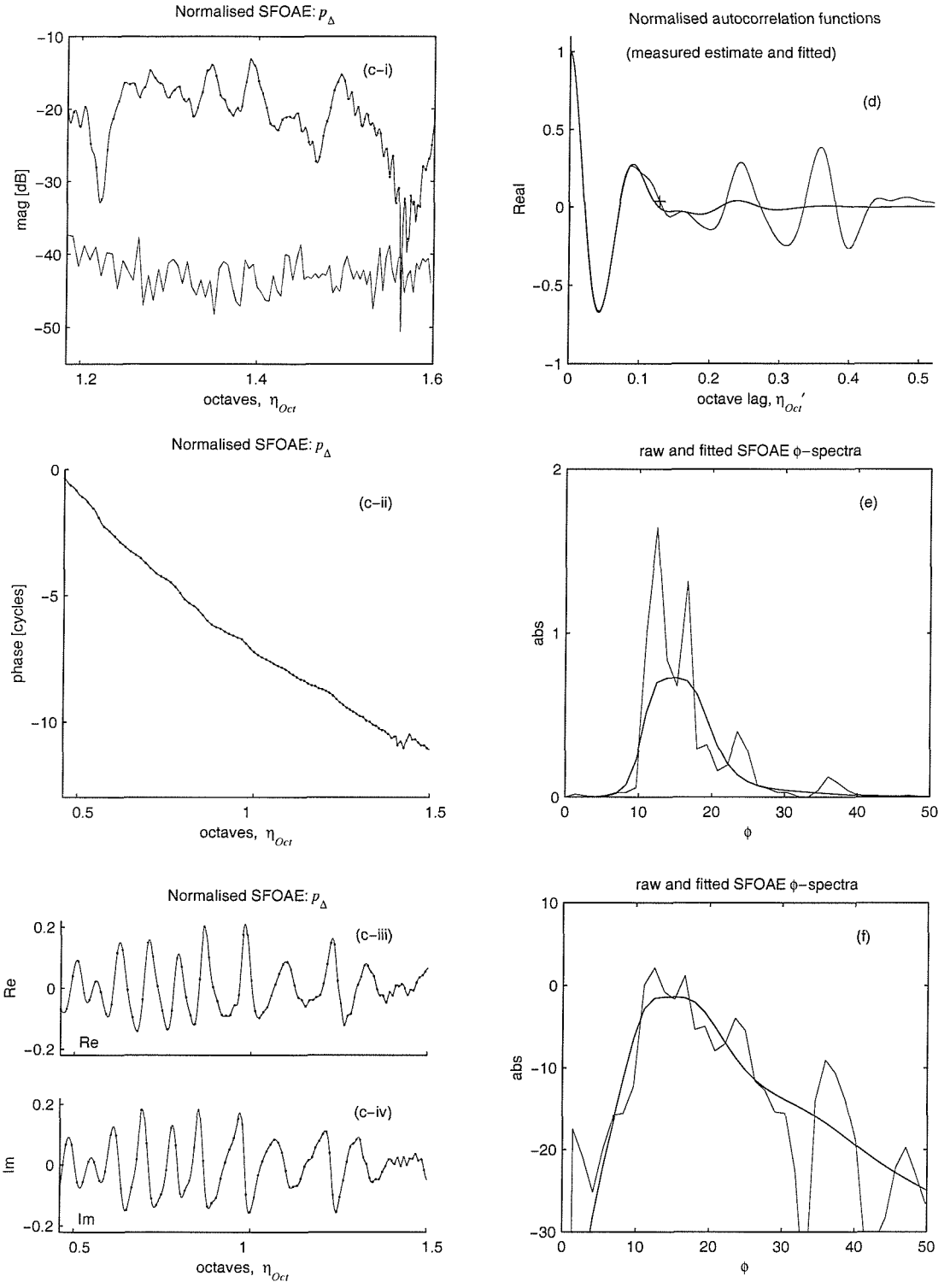


Fig 9.1 continued. Panels (c-i)-(c-iv): magnitude, phase, real and imaginary parts respectively, of the normalised SFOAE pressure, $p_\Delta \equiv p_{SF} / p_{EC:\,R=0}$ plotted against the logarithmic frequency, $\eta_{Oct} \equiv \log_2(f_1 / f_{Ref})$; $f_{Ref} = 1$ kHz, which has been resampled at equal η_{Oct} intervals (thin line is (c-i) indicates the noise floor). Panel (d) real part of the estimated measured normalised autocorrelation function of the normalised SFOAE (thin line), and the corresponding 4-parameter fit (thick line). Panel (e) and (f) the raw measured SFOAE ϕ-spectrum (thin line) and the fitted SFOAE ϕ-spectrum (thick line) on linear and logarithmic ordinates respectively.

Note that in fig 9.1a (i)-(iv) the maximally suppressed spectrum, $\boldsymbol{p}_{EC:R=0}(f)$, resembles a frequency-smoothed version of the unsuppressed spectrum, $\boldsymbol{p}_{EC}(f)$, as expected (section 2.9). This also suggests that the nonlinear SFOAE components predicted by Talmadge *et al.* (2000) are insignificant over the measured frequency range. If present, these components would differ in the suppressed and unsuppressed cases, thus leading to a deviation of the suppressed spectrum from the frequency-smoothed, unsuppressed spectrum. Since the predicted nonlinear components show no fine structure (i.e., they vary only very slowly with frequency) this deviation would be easily distinguished from SFOAE components arising from TW scattering.

Step 2

In step 2, the logarithmic frequency variable, η_{Oct}, is calculated, such that it equals the number of octaves above a reference frequency of 1 kHz.

$$\eta_{oct} \equiv \log_2\left(f_1/f_{Ref}\right); \qquad f_{Ref} = 1 \text{ kHz} \qquad [9.2]$$

Then the normalised SFOAE, $\boldsymbol{p}_{\Delta}(f_1)$, is resampled from the 91 values at 16-Hz intervals, to 1456 values at 1-Hz intervals (using a low-pass interpolation algorithm). It is then further interpolated (using linear interpolation) to give 512 values, denoted $\boldsymbol{p}_{\Delta}(n)$, at equal η_{Oct} intervals of $\Delta\eta_{Oct}$, which are at approximately 1/512 octaves. For simplicity of notation, the independent variable n is simply the index, 0, 1, 2, ...511.

Step 3

An estimate of the normalised autocorrelation function is then calculated. The estimated autocorrelation function could be calculated from:

$$\tilde{\mathrm{R}}_{p_\Delta p_\Delta}(m) \equiv \frac{1}{N}\sum_{n=0}^{N-m-1} \boldsymbol{p}_{\Delta}{}^{*}(n)\boldsymbol{p}_{\Delta}(n+m)$$
$$m = 0,1,2...N-1 \qquad [9.3]$$
$$n = 0,\ 1,\ 2...N-1$$
$$N = 512$$

where * denotes the complex conjugate. (For simplicity, only positive lag values are calculated.) However, rather than using equation [9.3], it has been found that a better estimate can be obtained by making use of the analytic properties of the signal, previously discussed in section 6.5. Briefly, for any given frequency sweep, Zweig and Shera's theory

predicts that the real and imaginary parts of $\boldsymbol{p}_\Delta(n)$ will not be independent, but will instead be related by the Hilbert transform. This is somewhat analogous to the relationship that holds for any frequency response function which corresponds to a causal (i.e., one-sided) impulse response function. Recall from section 2.13 that causality in the time domain leads to a frequency response whose real and imaginary parts form a Hilbert transform pair. In our case, it is the real and imaginary parts of the η-series that form the Hilbert transform pair, and the ϕ-spectrum which is one-sided.

This relationship is used to reject noise from the $\boldsymbol{p}_\Delta(n)$ signal, since the noise does not show the same analytic properties shown by the signal. This noise rejection occurs automatically when the following estimate of the autocorrelation function is used:

$$\tilde{R}_{p_{\Delta:\mathrm{Re}}\tilde{p}_{\Delta:\mathrm{Re}}}(m) \equiv \frac{2}{N}\sum_{n=0}^{N-m-1} p_{\Delta:\mathrm{Re}}(n)\tilde{p}_{\Delta:\mathrm{Re}}(n+m)$$

where

$$p_{\Delta:\mathrm{Re}}(n) \equiv \mathrm{Re}\{\boldsymbol{p}_\Delta(n)\}$$
$$p_{\Delta:\mathrm{Im}}(n) \equiv \mathrm{Im}\{\boldsymbol{p}_\Delta(n)\} \qquad [9.4]$$
$$\tilde{p}_{\Delta:\mathrm{Re}}(n) \equiv \mathrm{Hilb}\{p_{\Delta:\mathrm{Im}}(n)\}$$
$$m = 0,1,2 \ldots N-1$$
$$n = 0,1,2 \ldots N-1$$
$$N = 512$$

and where Hilb denotes the Hilbert transform. Equation [9.4] estimates only the real part of the autocorrelation function in [9.3], but the imaginary part can then be calculated assuming analyticity. In [9.4], first the Hilbert transform of the imaginary part of $\boldsymbol{p}_\Delta(n)$ is used to give an estimate of the real part, $\tilde{p}_{\Delta:\mathrm{Re}}$. Then the cross-correlation is calculated between the actual real part of $\boldsymbol{p}_\Delta(n)$ and this estimate. The validity of this method is discussed in sections 9.3.

Step 4

The 4-parameter model is then applied to the estimate of the autocorrelation function, $\tilde{R}_{p_\Delta\tilde{p}_{\Delta:\mathrm{Re}}}(m)$, yielding the estimates of the four parameters: ϕ_C, ϕ_{BW}, α and β. The estimated power ϕ-spectral density is then reconstructed from these parameters. In addition to the four parameters, the poorness-of-fit is also calculated as the difference between the

fitted and measured autocorrelation functions, $\mathsf{K}_{p_{\Delta:\mathrm{Re}}p_{\Delta:\mathrm{Re}}}(m)$ and $\tilde{\mathsf{K}}_{p_{\Delta:\mathrm{Re}}\tilde{p}_{\Delta:\mathrm{Re}}}(m)$ respectively:

$$\varepsilon_{MS}^2 \equiv \frac{1}{M_{MAX}} \sum_{m=1}^{M_{MAX}} \left(\tilde{\mathsf{K}}_{p_{\Delta:\mathrm{Re}}\tilde{p}_{\Delta:\mathrm{Re}}}(m) - \mathsf{K}_{p_{\Delta:\mathrm{Re}}p_{\Delta:\mathrm{Re}}}(m)\right)^2 \qquad [9.5]$$

Here K denotes the normalised autocorrelation function :

$$\tilde{\mathsf{K}}_{p_{\Delta:\mathrm{Re}}\tilde{p}_{\Delta:\mathrm{Re}}}(m) \equiv \tilde{\mathsf{R}}_{p_{\Delta:\mathrm{Re}}\tilde{p}_{\Delta:\mathrm{Re}}}(m) / \tilde{\mathsf{R}}_{p_{\Delta:\mathrm{Re}}\tilde{p}_{\Delta:\mathrm{Re}}}(0) \qquad [9.6]$$

and M_{MAX} is the maximum η-lag value used in the fit (see section 6.6). The logarithmic form of the poorness-of-fit is also useful:

$$L_\varepsilon = 20\log_{10}(\varepsilon_{MS}) \qquad [9.7]$$

Note that it is the poorness-of-fit that is minimised during the iterative parameter fitting procedure. The quantity in [9.5] is then the final minimum achievable value of this poorness-of-fit.

9.2 Assessing the Effect of SOAEs

It is important to ensure that the measured SFOAEs frequency sweeps are not strongly influenced by the presence of SOAEs, which may become synchronised to the probe tone. To assess SOAEs, measurements were made of the power spectrum of the ear canal pressure in the absence of any stimulus before and after each level series (section 8.5). From visual inspection of the spectrum, SOAEs were identified as any sharp peak in the spectrum with a bandwidth of 32 Hz or less, and which exceeded the surrounding background noise by at least 1 dB.

Of the 20 subjects tested, 9 subjects showed one or more SOAEs in the frequency range of interest. It is argued here that an SOAE is only likely to influence any given SFOAE frequency sweep if the acoustic power emitted spontaneously is of similar magnitude to that evoked by the probe tone. If the SOAE is very weak, then, according to Zweig and Shera's theory, the measured SFOAE will still be dominated by the mechanisms modelled in section 4. The approach taken here is to give each frequency sweep a rating, referred to as the 'spontaneous-to-evoked OAE rating' (or S/EOAE rating), which gives a crude indication of the likelihood that the frequency sweep will be significantly influenced by an SOAE. The rating is used in later sections of the analysis as a means of classifying the SFOAE frequency sweeps, and is calculated as follows. For each SOAE of frequency f_{SOAE} the level of the SOAE, $\mathrm{L}_{\mathrm{SOAE}}$ (in dB SPL), is compared to the level of the SFOAE

pressure, $L_{SF}(f_{SOAE})$, measured at the SOAE frequency, where $L_{SF}(f_{SOAE}) \equiv 10\log_{10}[\,|\boldsymbol{p}_{SF}(f_{SOAE})|^2/p_{Ref}^2]$, and $p_{Ref} = 20$ μPa. Where a subject has more than one SOAE, the emission for which $L_{SOAE} - L_{SF}(f_{SOAE})$ is the smallest is chosen.

The S/EOAE rating for each frequency sweep is then defined as:

'zero' if no SOAE could be measured

'low' if $L_{SOAE} + 6$ dB $< L_{SF}(f_{SOAE})$

'high' if $L_{SOAE} + 6$ dB $\geq L_{SF}(f_{SOAE})$

This is illustrated in fig. 9.2 for two subjects. Thus, for a sweep with a 'low' S/EOAE rating, the acoustic pressure of the emission (in the 16-Hz measurement band containing f_{SOAE}) is at least doubled in amplitude by the presentation of the probe tone.

In fig. 9.2 (a) it is clear that the peaks in the SFOAE magnitude coincide with the frequencies of the SOAEs, even for frequency sweeps with a 'low' S/EOAE rating (and even in some cases where the SFOAE level exceeds the corresponding SOAE level by over 10 dB). This fact does not, however, imply that the SOAE is influencing the SFOAE. Instead, according to Shera and Zweig's theory, the two phenomena are closely related with SOAEs being interpreted as self-sustaining SFOAEs in which the product of the cochlear apical and basal reflectances equals 1. However, on presenting an evoking tone of sufficient magnitude, self-suppression of the TW will rapidly reduce the apical reflectance, and thus the measured emission is essentially no different from a pure SFOAE. The reason for the coincidence of the peaks in the SFOAE with the frequencies of the SOAEs is that multiple reflections in the cochlea are involved in both phenomena.

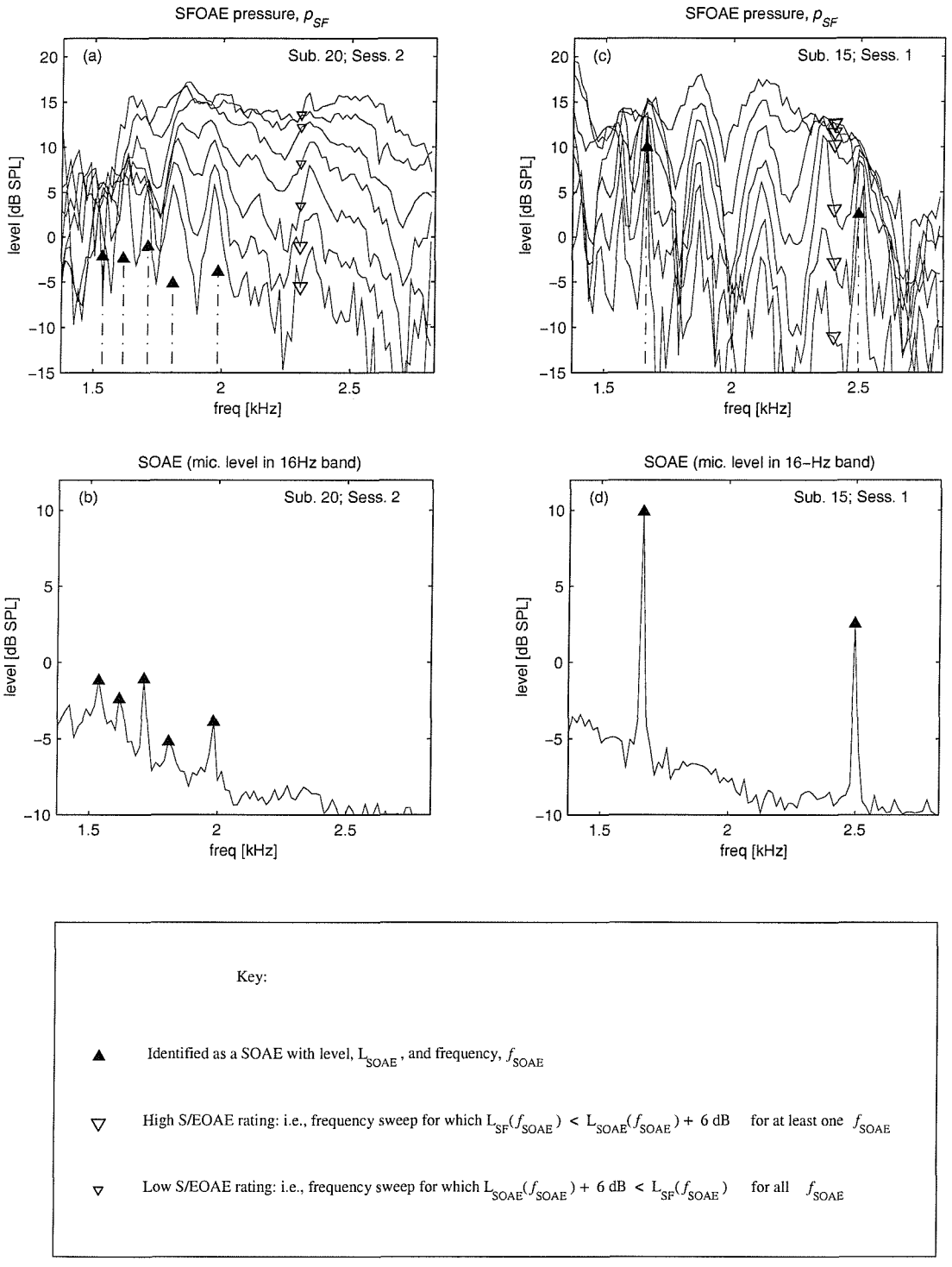


Fig 9.2 SFOAE self-suppression frequency sweeps for subjects with SOAEs. The figure also shows the classification of SFOAEs frequency sweeps according to whether SOAEs are expected to significantly influence the frequency sweep. Each SFOAE frequency sweep is given a 'S/EOAE strength' rating based on the strength of the SOAE relative to the SFOAE, measured at that frequency where the SOAE pressure is greatest, relative to the SFOAE pressure. A 'low' rating indicates that the measured frequency sweep is deemed likely to be dominated by the evoked OAE. A 'high' rating indicates that the measured frequency sweep is deemed likely to be influenced by an SOAE, at the SOAE frequency. (a) shows the SFOAE pressure, p_{SF}, over the seven frequency sweeps for the self-suppression experiment, for the subject with the SOAEs shown in (b) (subject 20, session 2). The S/EOAE strength is classed as 'high' for the two lower frequency sweeps, and 'low' for the five upper frequency sweeps. (c) and (d) show the corresponding data for a different subject (subject 15, session 1). For this subject, the S/EOAE strength is classed as 'high' for all seven frequency sweeps.

9.3 Estimation of SNR and Measurement Repeatability

Before presenting the measured data, the issue of measurement reliability is addressed. Frequency sweeps that are excessively contaminated with noise are excluded from the analysis since in these cases the 4-parameter model is unable to perform reliably. However, when excluding data, care must be taken not to distort the results (i.e., not to exclude data simply because they do not support the theory under test). It was decided to exclude frequency sweeps purely on the basis of an estimate of the SNR for that sweep.

Three methods of estimating the SNR for the measurements of the SFOAE have been used:

- SNR estimate 1: based on the additive noise present in adjacent frequency bands.
- SNR estimate 2: based on session-to-session repeatability.
- SNR estimate 3: based on the analytic properties of the SFOAE.

Of these, SNR estimate 1 was chosen as the reference SNR used to exclude frequency sweeps. The other two estimates are used to give additional indications of the quality of the measurements.

SNR estimate 1, denoted SNR_1, is calculated from the estimate of additive noise, defined in section 8.8.3, by averaging it over the frequency sweep. Thus there is one value of SNR_1 for each frequency sweep.

$$
\begin{aligned}
&\mathrm{SNR}_1 \equiv 10\log_{10}\left(\frac{S_1^2 - N_1^2}{N_1^2}\right) \\
&S_1^2 \equiv \frac{1}{(n_2 - n_1)}\sum_{n=n_1}^{n_2}\left|\boldsymbol{p}_{SF}(n)\right|^2 \\
&N_1^2 \equiv \frac{1}{(n_2 - n_1)}\sum_{n=n_1}^{n_2} n_{SF:Add}^2(n) \\
&n_1\Delta f = 1376\ \mathrm{Hz} \\
&n_2\Delta f = 2816\ \mathrm{Hz} \\
&\Delta f = 16\ \mathrm{Hz}
\end{aligned}
\qquad [9.8]
$$

The reason that the noise appears in the numerator in [9.8] is that the value of S_1^2 gives an estimate of the signal power plus the noise power.

SNR estimate 2, denoted SNR_2, is calculated from the correlation coefficient between the SFOAE frequency sweeps measured in two sessions. Thus there is one value of SNR_2 for each pair of frequency sweeps. The SNR is obtained directly from the correlation coefficient, ρ_{12}, by assuming (for simplicity) that any lack of correlation is due entirely to

additive noise. With this assumption, the following relationship between SNR and correlation coefficient is easily derived:

$$\mathrm{SNR}_2 \equiv 10\log_{10}\left(\frac{\rho_{12}}{1-\rho_{12}}\right)$$

$$\rho_{12} \equiv |\rho_{12}|$$

$$\rho_{12} \equiv \frac{1}{(n_2 - n_1)} \sum_{n=n_1}^{n_2} \boldsymbol{p}^*_{SF:Session\ 1}(n)\boldsymbol{p}_{SF:Session\ 2}(n) \qquad [9.9]$$

$$n_1 \Delta f = 1376\ \mathrm{Hz}$$

$$n_2 \Delta f = 2816\ \mathrm{Hz}$$

$$\Delta f = 16\ \mathrm{Hz}$$

where the * denotes the complex conjugate. Thus for example, a correlation coefficient of 0.5 gives $\mathrm{SNR}_2 = 0$ dB, whilst a correlation coefficient of 0.8 gives $\mathrm{SNR}_2 = 6$ dB.

SNR estimate 3, denoted SNR_3, is calculated from the correlation coefficient between the real part of $\boldsymbol{p}_\Delta$, denoted $p_{\Delta:\mathrm{Re}}$, and an estimate of $p_{\Delta:\mathrm{Re}}$, obtained from the Hilbert transform of the imaginary part of p_Δ. This estimate of $p_{\Delta:\mathrm{Re}}$ is denoted $\tilde{p}_{\Delta:\mathrm{Re}}$. Thus:

$$\mathrm{SNR}_3 \equiv 10\log_{10}\left(\frac{\rho_{p_{\Delta:\mathrm{Re}}\tilde{p}_{\Delta:\mathrm{Re}}}}{1-\rho_{p_{\Delta:\mathrm{Re}}\tilde{p}_{\Delta:\mathrm{Re}}}}\right)$$

$$\rho_{p_{\Delta:\mathrm{Re}}\tilde{p}_{\Delta:\mathrm{Re}}} \equiv \frac{\sigma^2_{p_{\Delta:\mathrm{Re}}\tilde{p}_{\Delta:\mathrm{Re}}}}{\sigma_{p_{\Delta:\mathrm{Re}}p_{\Delta:\mathrm{Re}}}\sigma_{\tilde{p}_{\Delta:\mathrm{Re}}\tilde{p}_{\Delta:\mathrm{Re}}}}$$

$$\sigma^2_{p_{\Delta:\mathrm{Re}}\tilde{p}_{\Delta:\mathrm{Re}}} \equiv \frac{1}{N}\sum_{n=0}^{N-1} p_{\Delta:\mathrm{Re}}(n)\tilde{p}_{\Delta:\mathrm{Re}}(n)$$

$$\sigma^2_{p_{\Delta:\mathrm{Re}}p_{\Delta:\mathrm{Re}}} \equiv \frac{1}{N}\sum_{n=0}^{N-1} p_{\Delta:\mathrm{Re}}(n)p_{\Delta:\mathrm{Re}}(n)$$

$$\sigma^2_{\tilde{p}_{\Delta:\mathrm{Re}}\tilde{p}_{\Delta:\mathrm{Re}}} \equiv \frac{1}{N}\sum_{n=0}^{N-1} \tilde{p}_{\Delta:\mathrm{Re}}(n)\tilde{p}_{\Delta:\mathrm{Re}}(n) \qquad [9.10]$$

where

$$p_{\Delta:\mathrm{Re}}(n) \equiv \mathrm{Re}\{\boldsymbol{p}_\Delta(n)\}$$

$$p_{\Delta:\mathrm{Im}}(n) \equiv \mathrm{Im}\{\boldsymbol{p}_\Delta(n)\}$$

$$\tilde{p}_{\Delta:\mathrm{Re}}(n) \equiv \mathrm{Hilb}\{p_{\Delta:\mathrm{Im}}(n)\}$$

$$m = 0,1,2 \ldots N-1$$

$$n = 0,\ 1,\ 2 \ldots N-1$$

$$N = 512$$

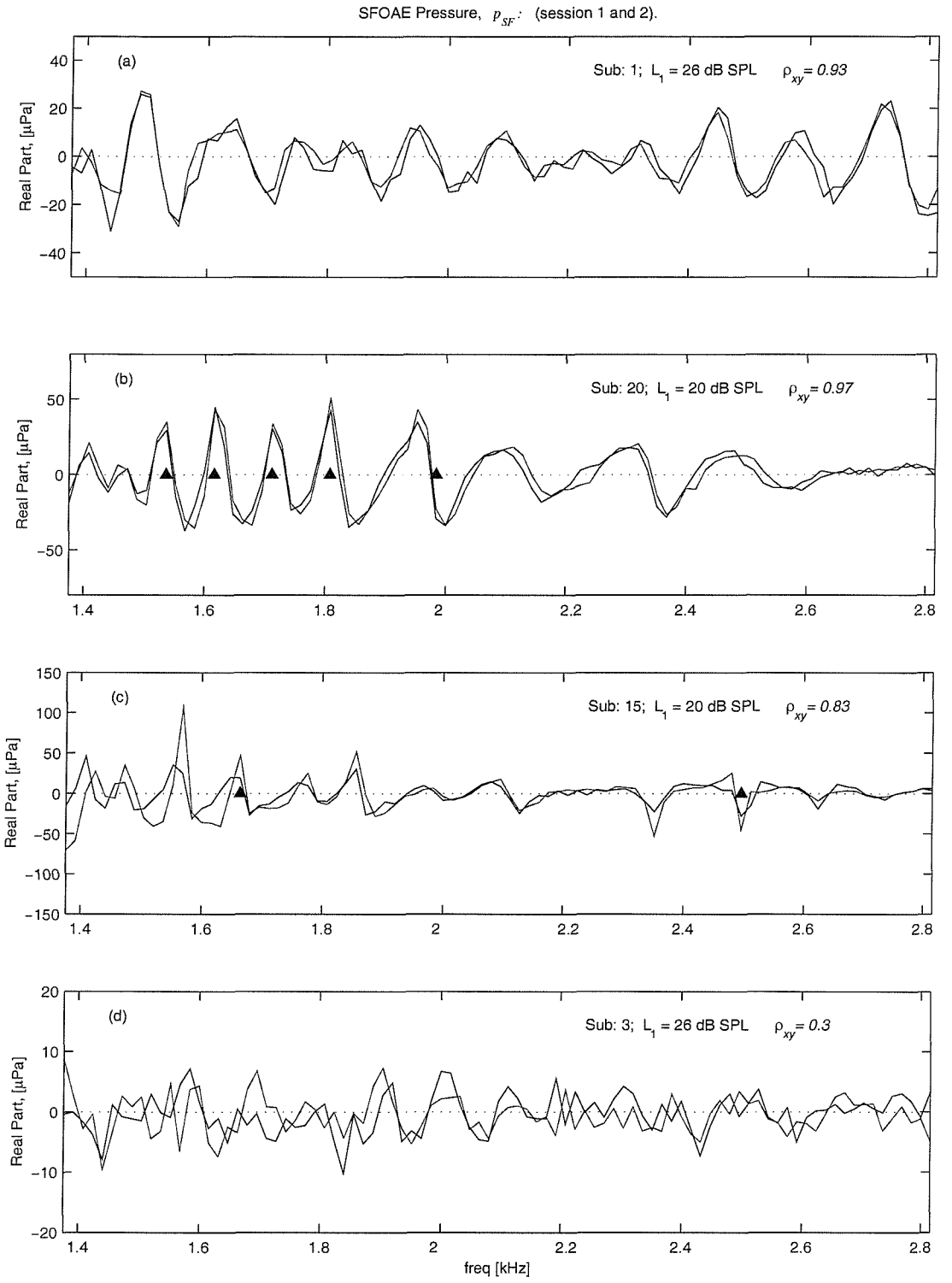


Fig 9.3 Session-to-session repeatability for four subjects. The real part of the SFOAE pressure for sessions 1 and 2 are overlaid for four subjects. The results are taken from the self-suppression experiment, from subjects with: (a) moderate SFOAEs, but no detectable SOAE (subject 1; L_1 = 26 dB SPL). (b) strong SFOAEs, and with both 'low' and 'high' S/EOAE strength ratings (subject 20, L_1 = 20 dB SPL) (c) strong SFOAEs and with 'high' S/EOAE strength ratings throughout (subject 15; L_1 = 20 dB SPL) (d) weak SFOAE and with no detectable SOAE (subject 3, L_1 = 26 dB SPL). Solid triangles indicate the location of SOAEs.

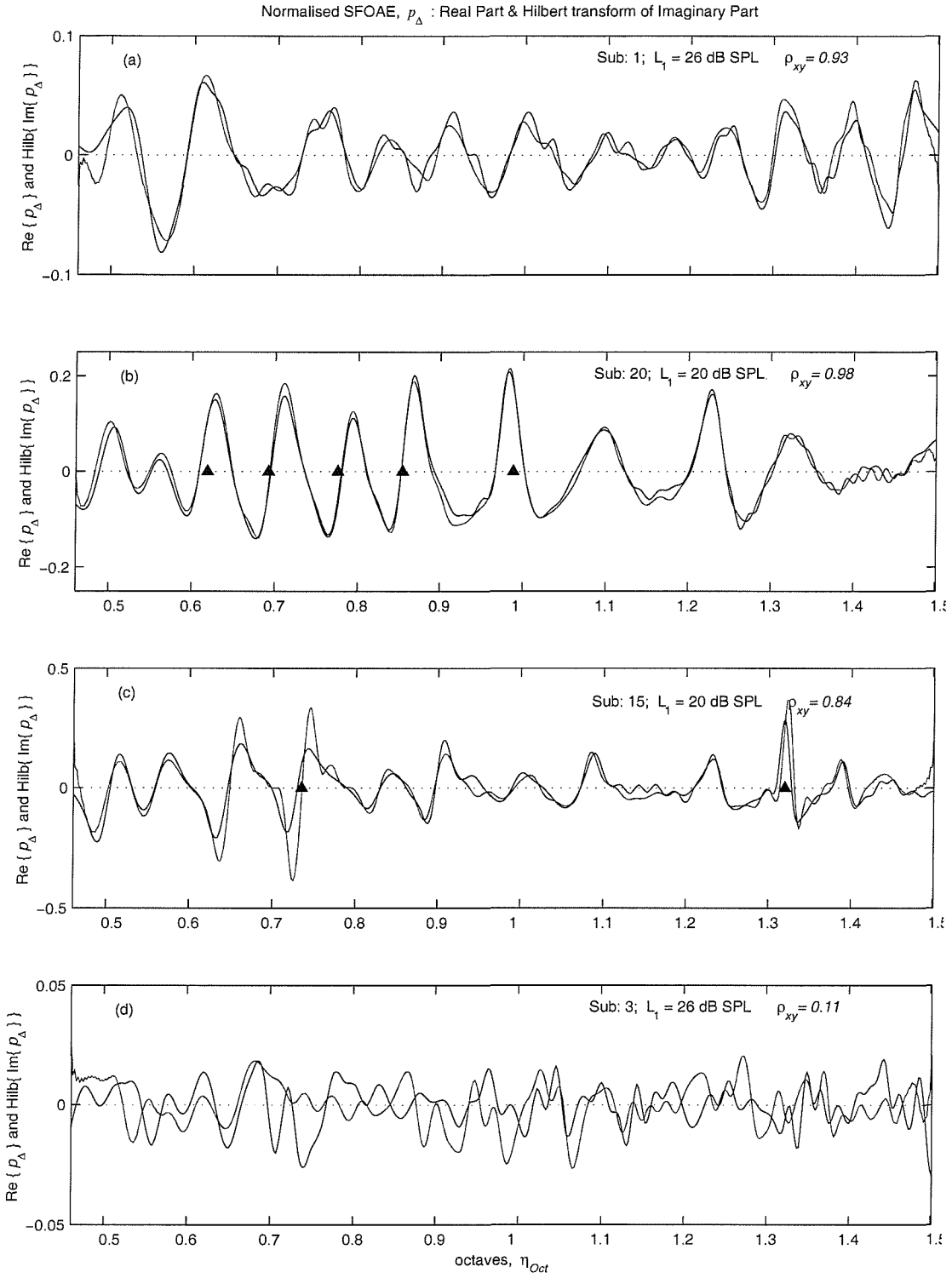


Fig 9.4 Analyticity of the SFOAE frequency sweep for four subjects. The real part of the normalised SFOAE is overlaid on top of the Hilbert transform of the imaginary part for four subjects. The results are for the self-suppression experiment, from subjects with (a) moderate SFOAEs, but no detectable SOAE (subject 1; L_1 = 26 dB SPL). (b) strong SFOAEs, and with both 'low' and 'high' S/EOAE strength ratings (subject 20, L_1 = 20 dB SPL) (c) strong SFOAEs and with 'high' S/EOAE strength ratings throughout (subject 15; L_1 = 20 dB SPL) (d) weak SFOAE and with no detectable SOAE (subject 3, L_1 = 26 dB SPL). Solid triangles indicate the location of SOAEs.

To illustrate the calculation of SNR_2, fig. 9.3 shows example comparisons of the frequency sweep for sessions 1 with that for session 2, for four subjects. The first subject had moderate SFOAEs, but no detectable SOAEs (subject 1), the second had strong SFOAEs and moderate SOAEs (subjects 20), the third had strong SFOAEs and strong SOAEs (subject 15), and the fourth subject had weak SFOAEs, and no SOAEs. The frequency sweeps from the first three subjects show good correlation between sessions. The frequency sweep from the fourth shows poor correlation, leading to a low estimate for SNR_2.

To illustrate the calculation of SNR_3, fig. 9.4 shows examples of the signal analyticity for the same frequency sweeps as in fig. 9.3. Note that similar values of the correlation coefficients are obtained, leading to similar values of SNR_2 and SNR_3.

9.4 Criterion for Rejecting Frequency Sweeps

It was decided to reject all frequency sweeps for which the SNR_1 value fell below 6 dB. The reasons for choosing SNR_1 over SNR_2 or SNR_3 are as follows. SNR_2 may be an underestimate of the true SNR, since some changes in the measured SFOAE frequency sweep will arise simply by session to session changes in the fit, or the middle ear properties. Were SNR_2 used as a criterion, relatively noise free frequency sweeps might be rejected unnecessarily. It was decided that SNR_3 should not be used as a rejection criterion, since its validity depends to some extent on the theory to be tested. This leaves SNR_1 as the preferred rejection criterion. SNR_1 does, however, have the disadvantage that it is insensitive to drift noise, which would be expected to show up in both SNR_2 and SNR_3. Therefore it is expected that SNR_1 will overestimate the true SNR.

Fig. 9.5a shows SNR_2 plotted against SNR_1 for each pair of frequency sweeps, for all the data. There are approximately 240 pairs of sweeps shown (20 subjects $\times$ 7 probe levels + 20 subjects $\times$ 5 suppressor levels). This shows that SNR_2 is typically about 5 dB lower than SNR_1, as expected for the reasons already discussed. Similarly, SNR_3 is typically 3 dB below SNR_1 (fig. 9.5b). Rejected frequency sweeps are indicated on the figure. Also shown is the poorness-of-fit, defined in equation [9.7]. This shows the expected result that the 4-parameter model performance improves with improvements in SNR. Note that values of the logarithmic poorness-of-fit greater than about –10 indicate a complete failure to converge to realistic parameters. The good correlation between all three estimates of SNR suggests that they are all reliable measures. Note also that although SNR_2 has not

been adopted as the basis for rejecting noisy data, inter-session comparisons of data are performed later in the analysis after estimates of the four parameters have been obtained for both sessions.

Table 9.1 shows a summary of the numbers of sweeps retained in the analysis. The target number of sweeps exceeds the actual number of sweeps as some of the sessions could not be completed due to time constraints.

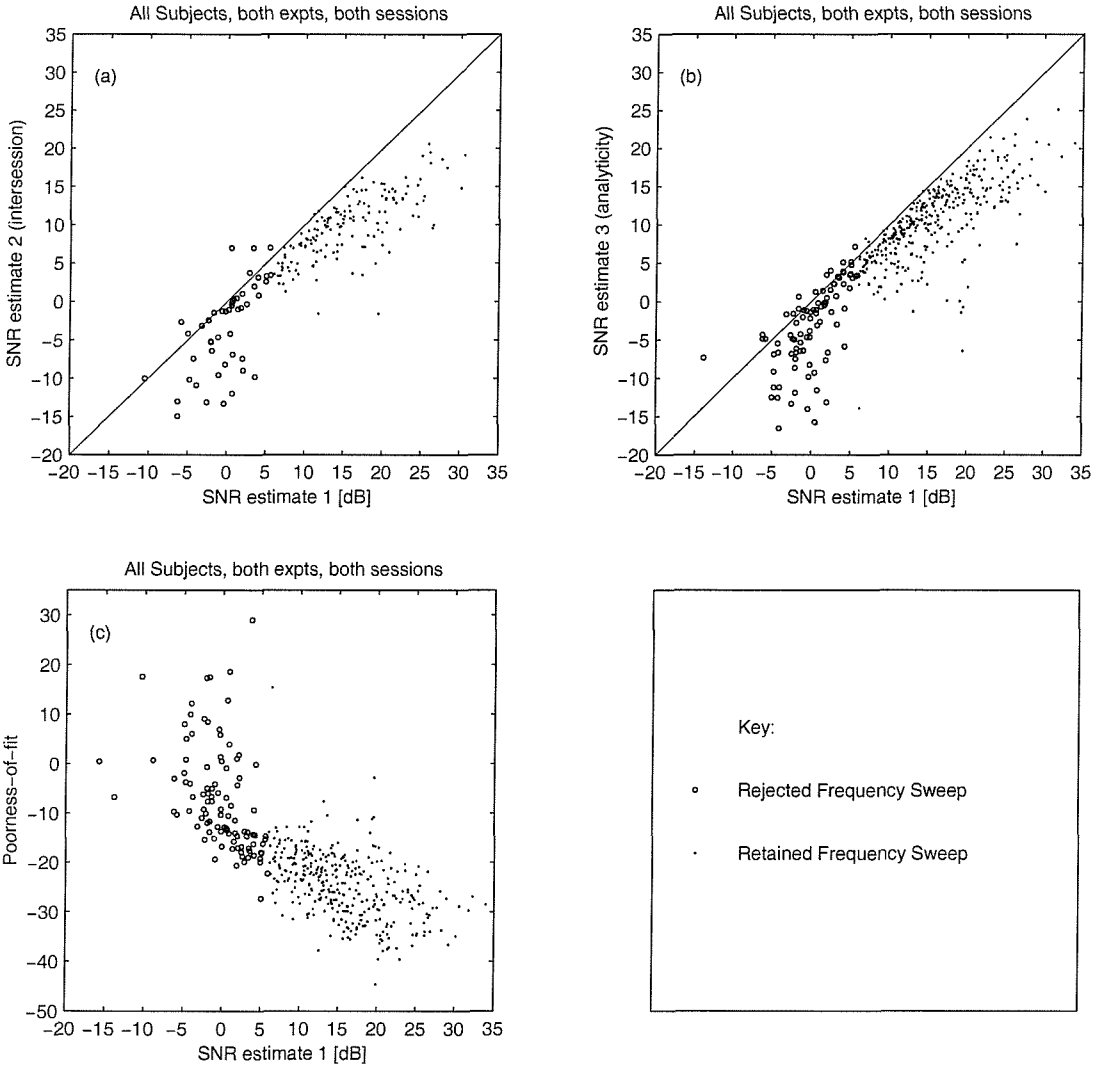


Fig 9.5 Measures of the quality of the data and the parametric fit. Three measures are shown, plotted against the SNR estimate 1 (the reference SNR estimate), which is obtained using the frequency average of the 'noise in adjacent bands' SNR estimate. Results are shown for all subjects, for both self- and two-tone suppression experiments and for both sessions. (a) shows the a second estimate of SNR, obtained by correlating the SFOAE frequency sweep from session 1 with that from session 2. There is one point for each pair of frequency sweeps. (b) shows a third estimate of the SNR, based on the assumption that the normalised SFOAE frequency sweep is analytic. It is derived from the correlation of the real part of the frequency sweep with the Hilbert transform of the imaginary part. There is one point per frequency sweep. (c) shows the logarithmic measure of the poorness-of-fit of the 4-parameter model for each frequency sweep. There is one point per frequency sweep. Also indicated are those frequency sweeps which are rejected because their SNR estimate 1 falls short of a set threshold value of 6 dB.

Table 9.1: Summary of frequency sweep rejections due to poor SNR

Numbers of frequency sweeps	Self-suppression	Two-tone suppression
Target	2×20×7=280	2×20×5=200
Actual	273	170
Rejected	53	47
Retained	220	123

Table 9.2 summarises the statistics of the S/EOAE ratings for the retained frequency sweeps.

Table 9.2: Breakdown of retained frequency sweeps by S/EOAE rating.

Numbers of frequency sweeps	Self-suppression	Two-tone suppression
Retained: Total	220	123
Retained with SOAEs absent	96	41
Retained with low S/EOAE rating	74	45
Retained with high S/EOAE rating	50	37

9.4.1 The Validity of Applying the 4-parameter Model to Experimental Data

The choice of the 2nd order Butterworth filter was based on predicted SFOAEs from cochlear models. As will be shown in later sections, where measured and fitted ϕ-spectra are compared, it appears that the 4-parameter model may also be applied successfully to experimental data. It might be thought that the poorness-of-fit could be used to quantify how well the measured data conformed to the 4-parameter model. However, the results in fig. 9.4 indicate that the poorness-of-fit is dominated by the SNR rather than by deficiencies in the stochastic model underlying the 4-parameter model.

9.5 Discussion of Results

Examples of the results are presented for four subjects who have been chosen to illustrate various features seen across all 20 subjects. The chosen subjects are:

1. subject 1, with moderately strong SFOAEs and with no detectable SOAEs;

2. subject 20, with strong SFOAEs and with moderate SOAEs, such that some frequency sweeps have a low S/EOAE rating, and others have a high S/EOAE rating.
3. subject 15, with strong SFOAEs and with strong SOAEs, such that all the frequency sweeps have a high S/EOAE rating.
4. subject 3 with weak SFOAEs and no detectable SOAEs. Several of the frequency sweeps are rejected due to poor SNR_1 estimates.

9.5.1 Qualitative Discussion of the Results for the Self-suppression Experiment

Fig. 9.6 shows the measured ear canal pressure in both the unsuppressed and maximally suppressed conditions for the self-suppression experiments for all four subjects. (Results from only one session are shown.) This shows the expected pattern of ripples, with a spacing that corresponds to the SFOAE periodicity. The ripple amplitude reduces as the probe level is increased.

Fig. 9.7 shows various results derived from the measurement of the complex SFOAE pressure for subject 1, plotted against the logarithmic frequency variable, η_{Oct}, for each probe level. Panel (a) shows that the magnitude of the SFOAE pressure, p_{SF}, increases with probe level, reaching a maximum of about 15 dB SPL. Panel (c) shows that the magnitude of the normalised SFOAE pressure, p_{Δ}, decreases with probe level, indicating the expected compressive nonlinearity. Panel (b) shows the SNR as a function of frequency, estimated using the noise in adjacent frequency bands, as in [8.1]. The SNR is typically well above 6 dB, indicating that the estimates of the SFOAE should be reliable. There is, however, a region at about $\eta_{Oct} = 1.2$ where the SNR is poor, and where therefore the estimated SFOAE is unreliable. Note that the SNR_1 estimate, used as a basis for rejecting frequency sweeps, can not be derived directly from the frequency dependent quantity plotted in panel (b), since SNR_1 is obtained by frequency averaging the signal and noise powers separately. Note also that, provided their SNR_1 estimate exceeds 6 dB, sweeps such as those in panel (b) are retained even when they contain regions where the frequency dependent SNR is very weak. The justification for this is that these parts of the signal are weak and have the least influence on the estimates given by the 4-parameter model.

The SFOAE periodicity does not show up in the magnitude plots in panels (a) and (c). Instead, the periodicity is most clearly seen as the ripple pattern in the real and imaginary parts of the normalised SFOAE, shown in panels (d) and (f). (Recall that a single frequency sweep with a constant periodicity of 1/15 would appear as a ripple pattern

with a peak to peak interval of about 0.1 octaves.) The periodicity is also seen in the phase of the normalised SFOAE, shown in panel (e). Here, a frequency sweep with a periodicity of 1/15 would appear as a straight line with slope of about 10 cycles per octave. Rather than being approximately linear, the actual phase measurements shown in panel (e) exhibit a slope which steepens with η_{Oct}. This result is in accordance with previous reports (e.g., Zweig and Shera, 1995, p. 2036). A possible explanation for this given by Zweig and Shera, is that the real cochlea departs from scaling symmetry due to a broadening of mechanical tuning at low characteristic frequencies. Note , however, that this variation in periodicity with frequency is not the subject of this thesis. Instead, the thesis concentrates on a form of frequency averaged periodicity, as extracted from the measurements by the 4-parameter model. It is variations in this periodicity with probe level that are of most interest. It is interesting to note that no obvious change in periodicity with probe level can be discerned from a visual inspection of the results shown in panels (d), (e) and (f). It is argued here that there is a significant change in periodicity with probe level, but that the representations in panels (d), (e) and (f) are not well suited to revealing this change. This change is, however, revealed by applying the 4-parameter model to the results as seen in fig. 9.8.

Fig. 9.8 shows the raw measured and the fitted ϕ-spectra corresponding to the η-series shown in fig. 9.7. The spectra are shown on both a linear and a dB ordinate. The linear scale is useful because the 4-parameter model performs a fit to the autocorrelation function, which is related (via the Fourier transform) to the linear spectrum. Therefore any features that appear insignificant in the raw linear spectrum will have no significant influence on the 4-parameter fit. In contrast, these feature will be magnified by the dB scale. The dB scale is useful in that it clearly shows the trends in the spectra that occur over the entire level series.

The most striking result is that the fitted spectra clearly show changes in ϕ-centre value, ϕ_C, with increasing probe level (panels (b) and (d)). This agrees with the predictions made by the cochlear models based on Zweig and Shera's theory (cf., fig 7.5). Variations in the bandwidth are less easy to judge simply by inspection. A quantitative discussion of this variation, and that of the other fitted parameters is given later. Two other points are worth noting. The first is that the variation in ϕ_C is not immediately obvious from the raw spectra. This is similar to the results obtained from the cochlear models (sections 5 and 6). The second is that the multiple reflection parameter, α, does not appear to vary smoothly, as seen by the erratic appearance of the second spectral lobe in panel (d) for nominal probe levels of 38 and 44 dB SPL. This is thought to be an inaccurate result, arising from the

difficulty in reliably estimating α. As will be discussed in more detail later, such errors in estimating α do not invalidate the estimates in the other parameters.

Figs. 9.9 and 9.10 show the SFOAE η-series and ϕ-spectra for subject 20. These results are in general qualitatively similar to those of subject 1, though the SFOAE level and SNR are greater. Two differences are worth noting. First, at the lower probe levels (14 – 26 dB SPL) there are clear peaks in the magnitude of the SFOAE (figs. 9.9a and c) coinciding with the presence of SOAEs. These peaks become less pronounced as the probe level increases. This is in accordance with the Zweig and Shera model, where multiple reflections become less significant as the magnitude of the apical reflectance is reduced. The second difference is that the phase curves (panel e) are more linear than for subject 20. In fact, in contrast to the trend seen in subject 1, at a nominal probe level of 14 dB SPL, the phase curve becomes less rather than more steep at higher frequencies. It was decided not to attempt to quantify this effect, or to study it further in this thesis. However, since this trend is opposite to that reported elsewhere (e.g., Zweig and Shera, 1995) it is perhaps worthy of a future investigation.

It can be seen in fig. 9.10 that, as for subject 1, the fitted ϕ-spectra show the trend of reducing ϕ_C with increasing probe level.

Figs. 9.11 and 9.12 show the SFOAE η-series and ϕ-spectra for subject 15. Again these results are in general qualitatively similar to those of subject 1. The main difference is that, even at high probe levels, there are clear peaks in the in the magnitude of the SFOAE (figs. 9.11a and c) coinciding with the presence of strong SOAEs. All the frequency sweeps here have high a S/EOAE rating, indicating that the frequency sweeps may be significantly affected by SOAEs. Despite the presence of strong SOAEs, the fitted ϕ-spectra (fig. 9.12) show the trend of reducing ϕ_C with increasing probe level, as for the previous subjects.

Figs. 9.13 and 9.14 show the SFOAE η-series and ϕ-spectra for subject 3, who had unusually weak SFOAEs. In these sweeps the SNR is poor, and therefore most of the sweeps are rejected. Only at the higher probe level (e.g., 44 and 50 dB SPL nominal) is there any sign of an SFOAE with the expected periodicity, as revealed by the phase curves in fig. 9.13 panel (e). The fitted ϕ-spectra are also erratic, showing little continuity between spectra from neighbouring probe levels. Since all the unusual looking spectra correspond to sweeps where the SNR was poor, it is likely that these are entirely unreliable results. In fact all these unusual looking spectra are eliminated, based on the low value of their SNR_1 estimate.

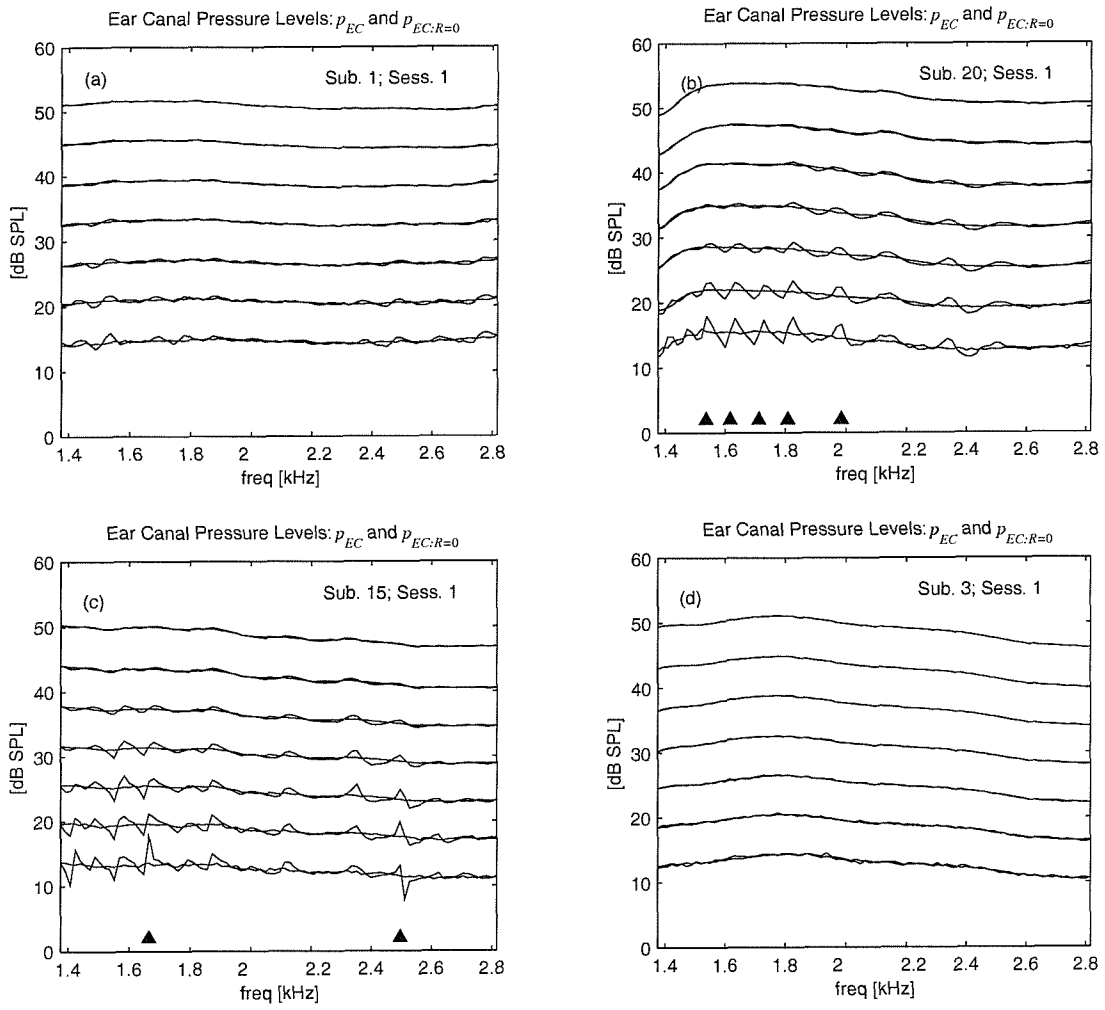


Fig 9.6 Ear canal sound pressure level for the self-suppression experiment for four subjects. The results are from subjects with (a) moderate SFOAEs, but no detectable SOAE (subject 1; session 1); (b) strong SFOAEs, and with both 'low' and 'high' S/EOAE strength ratings (subject 20, session 1); (c) strong SFOAEs and with 'high' S/EOAE strength ratings throughout (subject 15; session 1); (d) weak SFOAEs and with no detectable SOAEs (subject 3, session 1). The unsuppressed ear canal pressure, p_{EC} , appears as a rippled trace, whilst the suppressed ear canal pressure, $p_{EC:R=0}$, appears as a smooth trace. Solid triangles indicate any SOAE frequencies.

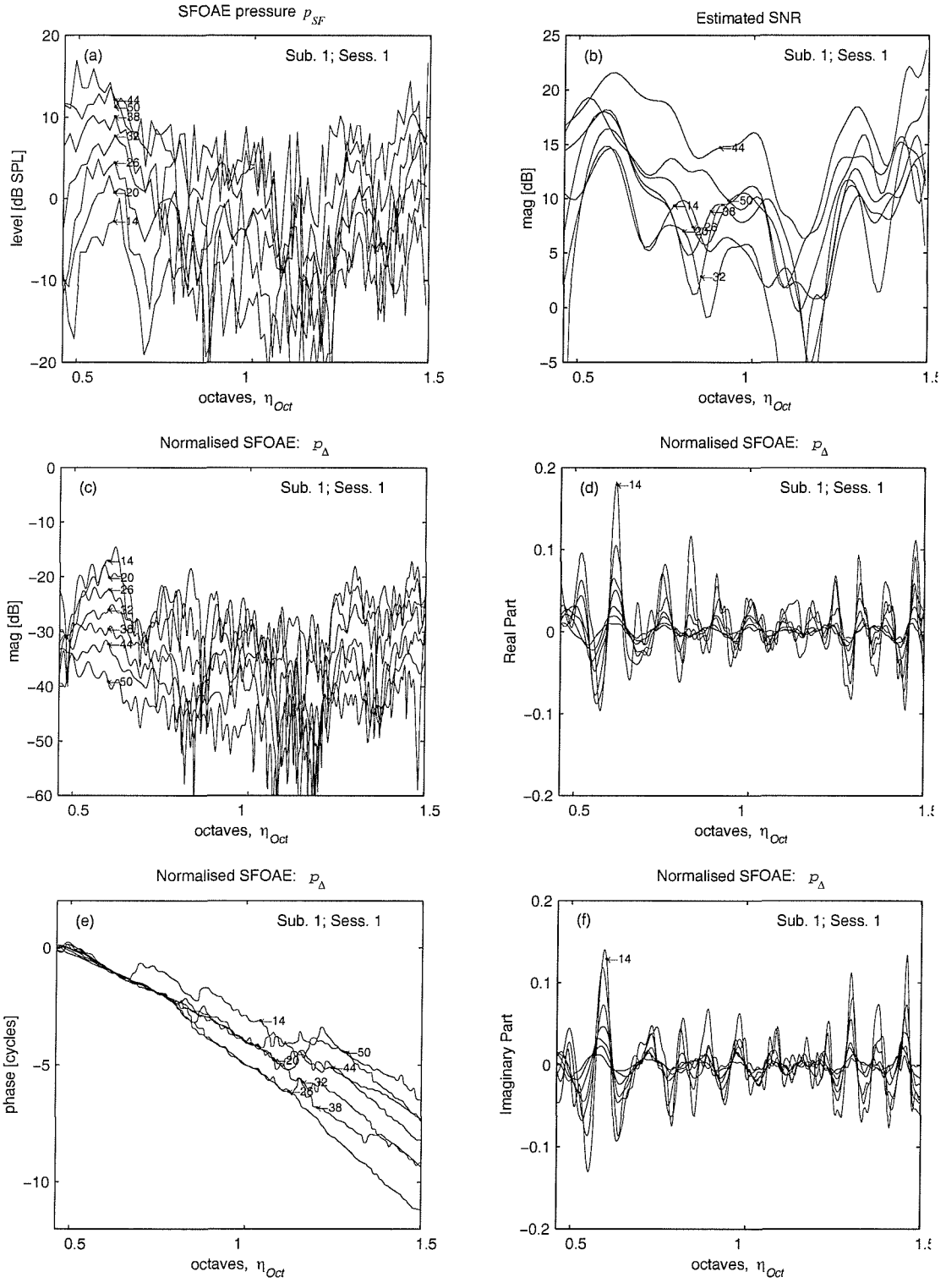


Fig 9.7 Measured SFOAE frequency sweep data for the self-suppression experiment, for a subject showing moderate SFOAEs, but with no detectable SOAEs (subject 1, session 1). Nominal probe tone levels are 14 to 52 dB SPL, indicated on curves. (a) SFOAE pressure level; (b) estimated SNR; (c) normalised SFOAE level (d) real part of normalised SFOAE; (e) phase of normalised SFOAE; (f) imaginary part of normalised SFOAE.

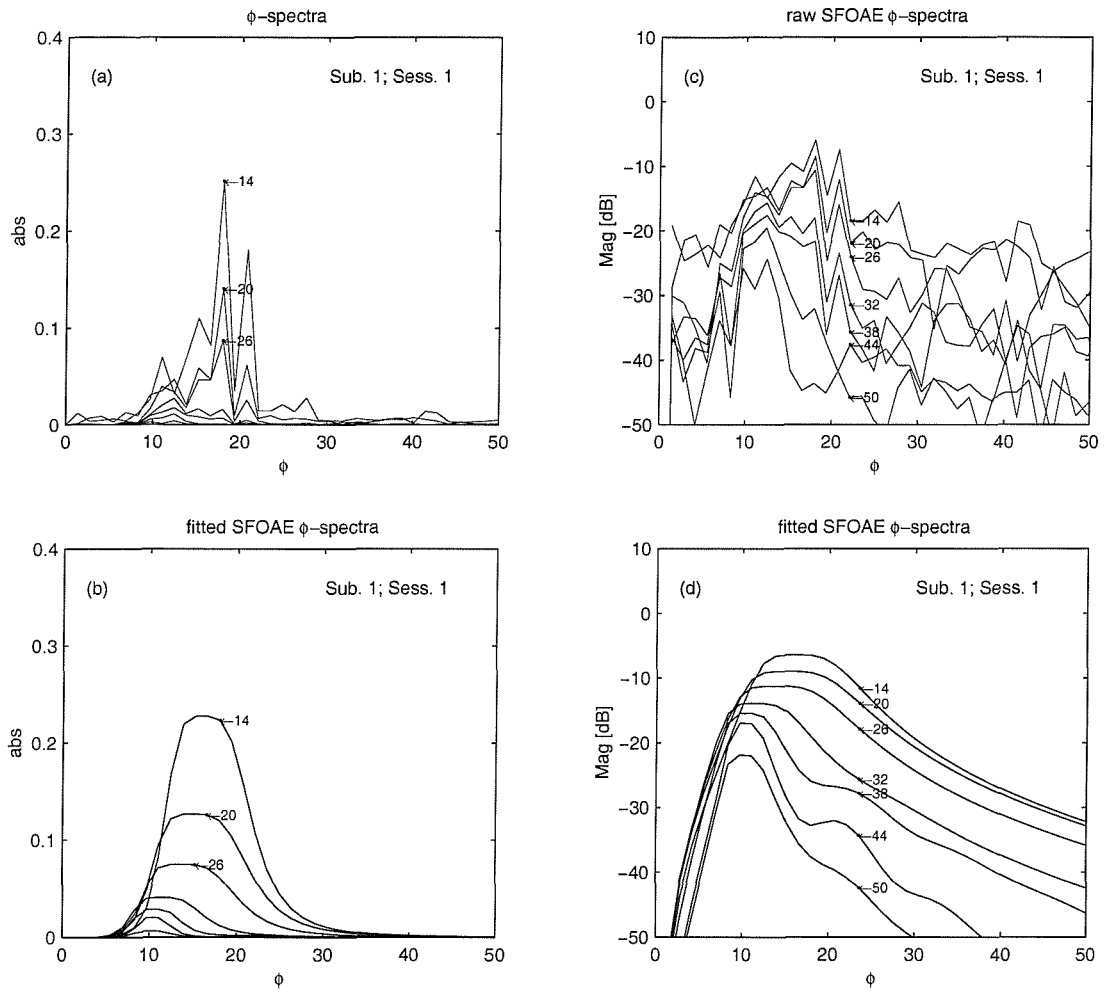


Fig 9.8 Raw and fitted SFOAE ϕ–spectra for the self-suppression experiment, for a subject showing moderate SFOAEs, but with no detectable SOAEs (subject 1, session 1). Nominal probe tone levels are 14 to 52 dB SPL, indicated on curves. (a) and (b) show raw and fitted ϕ–spectra plotted on a linear vertical scale. (c) and (d) show raw and fitted ϕ–spectra plotted on a dB scale.

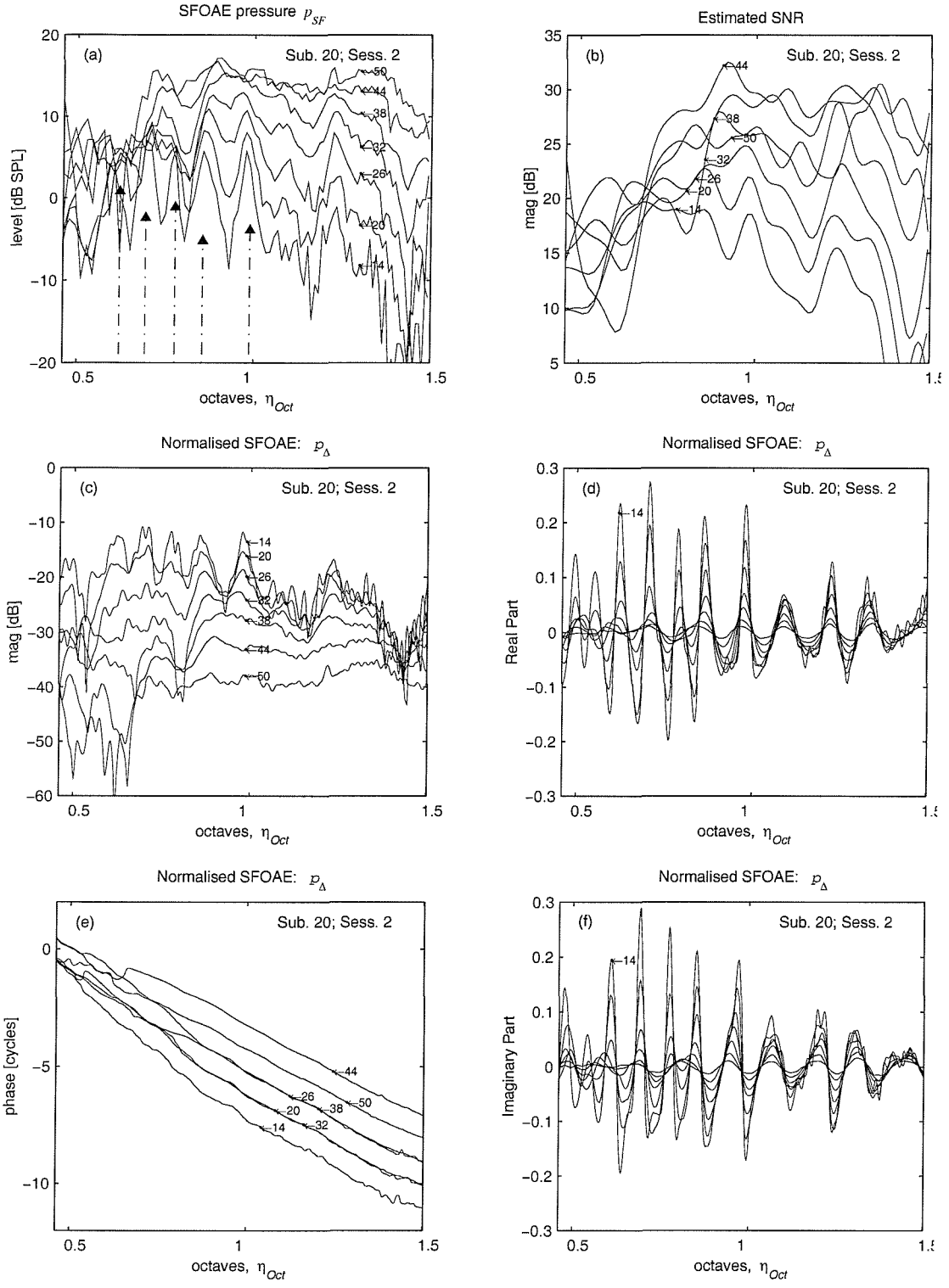


Fig 9.9 Measured SFOAE frequency sweep data for the self-suppression experiment, for a subject showing strong SFOAEs, and with both 'low' and 'high' S/EOAE strength ratings (subject 20, session 2). Nominal probe tone levels are 14 to 52 dB SPL, indicated on curves. (a) SFOAE pressure level; (b) estimated SNR; (c) normalised SFOAE level (d) real part of normalised SFOAE; (e) phase of normalised SFOAE; (f) imaginary part of normalised SFOAE.

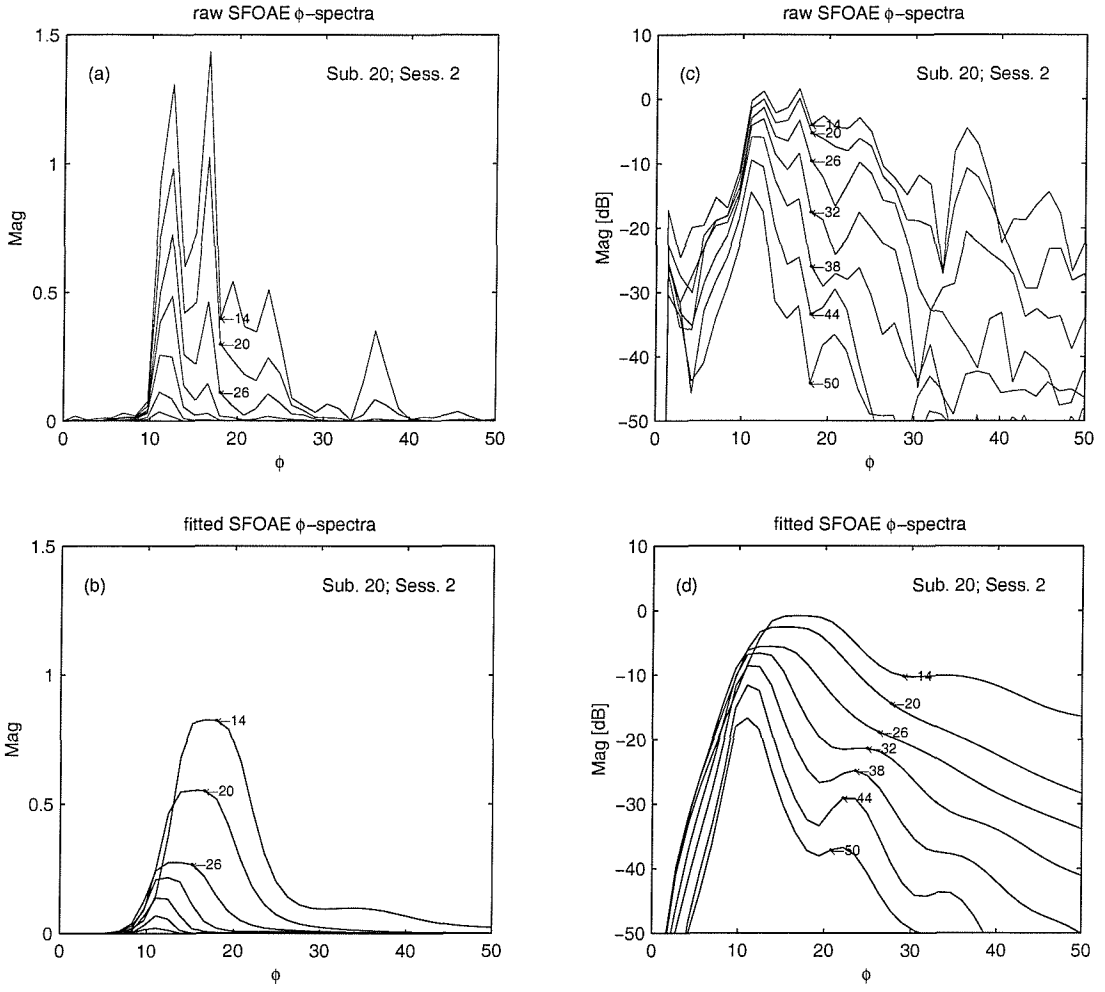


Fig 9.10 Raw and fitted SFOAE ϕ–spectra for the self-suppression experiment, for a subject showing strong SFOAEs, and with both 'low' and 'high' S/EOAE strength ratings (subject 20, session 2). Nominal probe tone levels are 14 to 52 dB SPL, indicated on curves. (a) and (b) show raw and fitted ϕ–spectra plotted on a linear vertical scale. (c) and (d) show raw and fitted ϕ–spectra plotted on a dB scale.

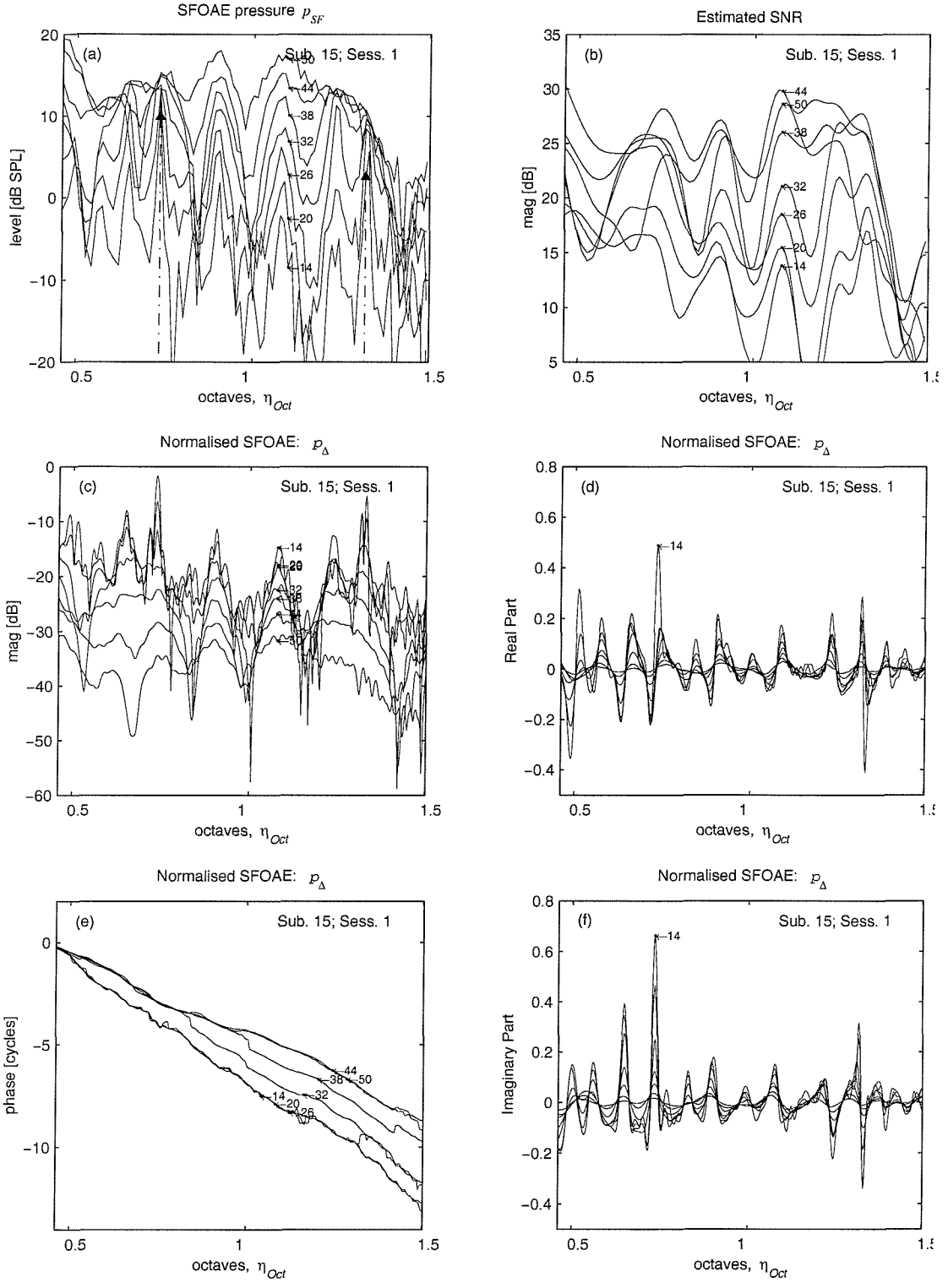


Fig 9.11 Measured SFOAE frequency sweep data for the self-suppression experiment, for a subject showing strong SFOAEs, and with 'high' S/EOAE strength ratings (subject 15, session 1). Nominal probe tone levels are 14 to 52 dB SPL, indicated on curves. (a) SFOAE pressure level; (b) estimated SNR; (c) normalised SFOAE level (d) real part of normalised SFOAE; (e) phase of normalised SFOAE; (f) imaginary part of normalised SFOAE.

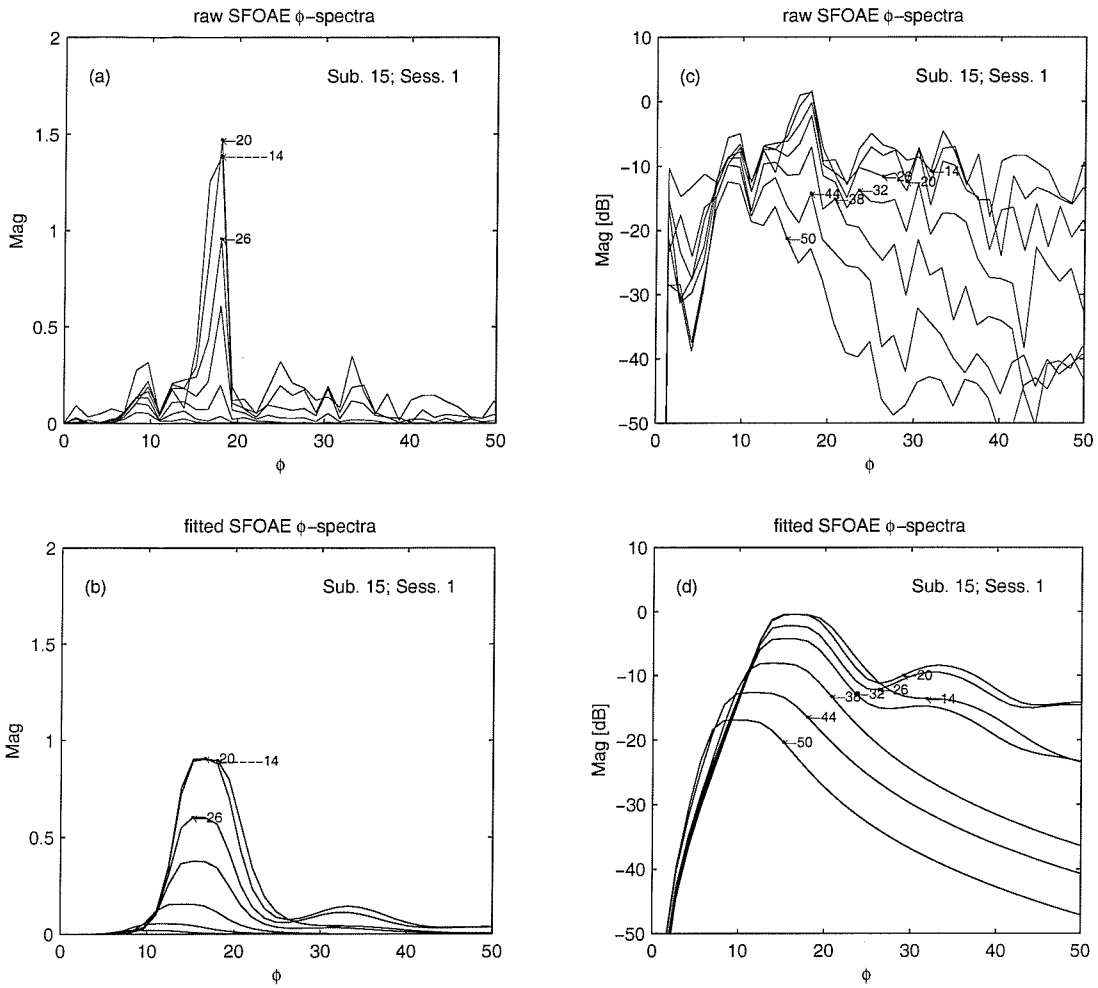


Fig 9.12 Raw and fitted SFOAE ϕ–spectra for the self-suppression experiment, for a subject showing strong SFOAEs, and with 'high' S/EOAE strength ratings (subject 15, session 1). Nominal probe tone levels are 14 to 52 dB SPL, indicated on curves. (a) and (b) show raw and fitted ϕ–spectra plotted on a linear vertical scale. (c) and (d) show raw and fitted ϕ–spectra plotted on a dB scale.

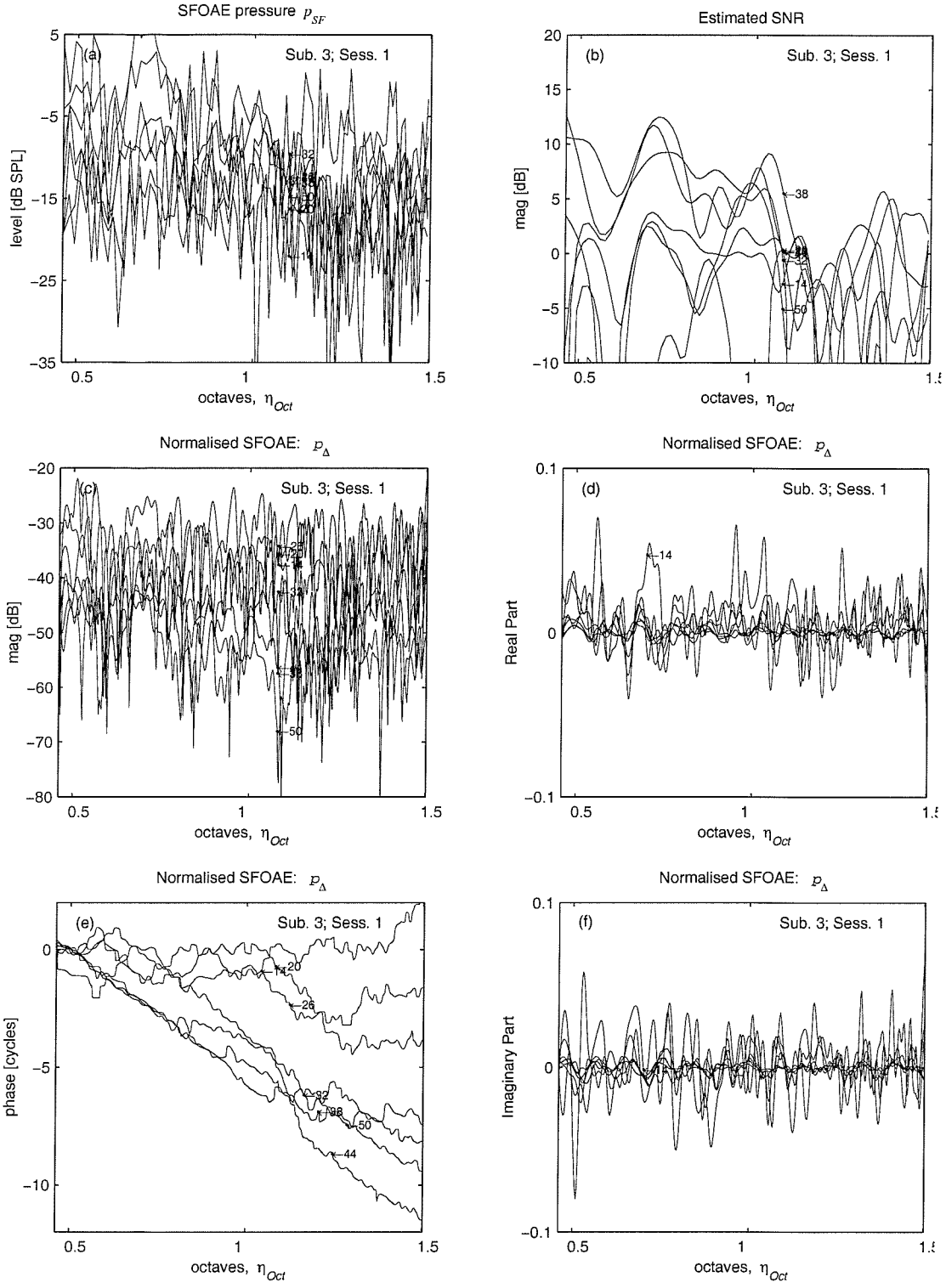


Fig 9.13 Measured SFOAE frequency sweep data for the self-suppression experiment, for a subject showing weak SFOAEs, and with no detectable SOAEs (subject 3, session 1). Nominal probe tone levels are 14 to 52 dB SPL, indicated on curves. (a) SFOAE pressure level; (b) estimated SNR; (c) normalised SFOAE level (d) real part of normalised SFOAE; (e) phase of normalised SFOAE; (f) imaginary part of normalised SFOAE.

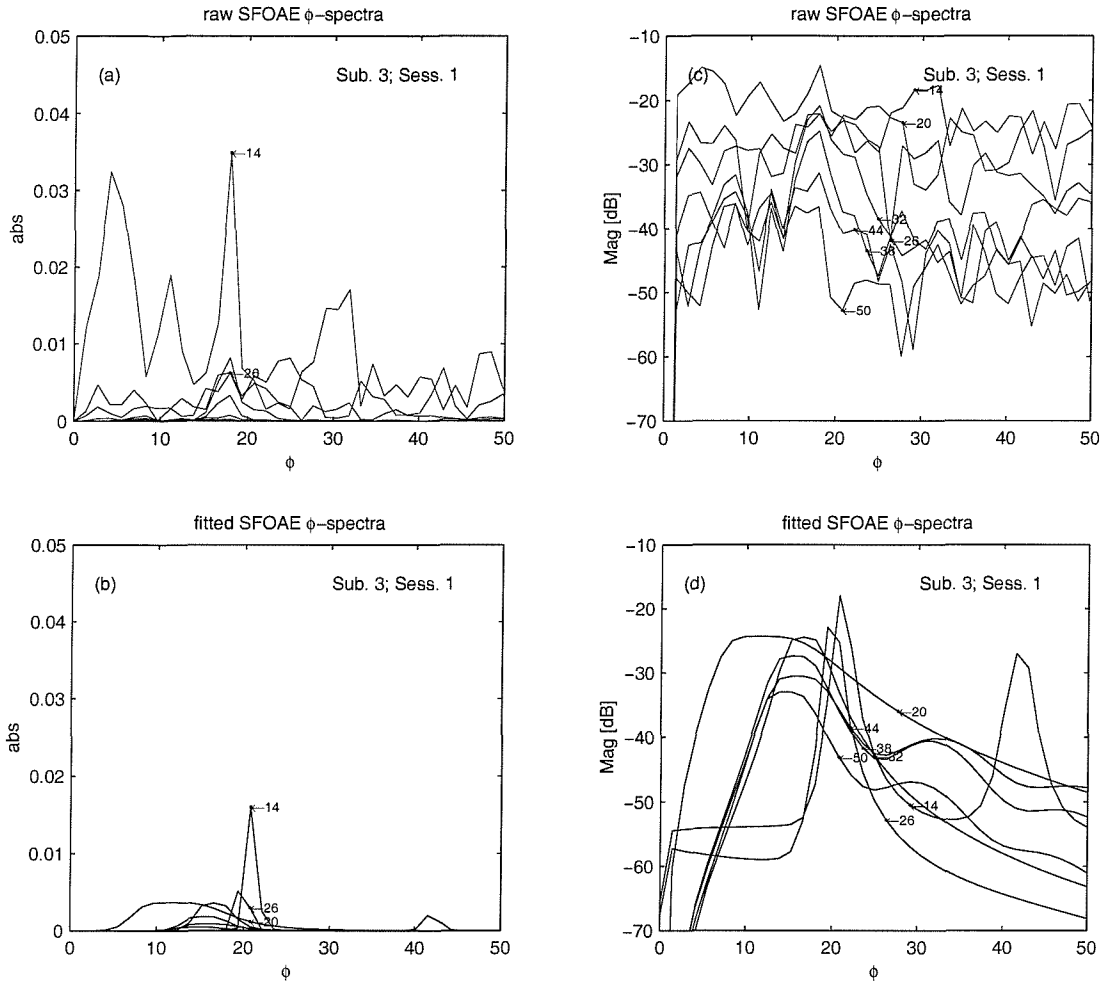


Fig 9.14 Raw and fitted SFOAE ϕ–spectra for the self-suppression experiment, for a subject showing weak SFOAEs, and with no detectable SOAEs (subject 3, session 1). Nominal probe tone levels are 14 to 52 dB SPL, indicated on curves. (a) and (b) show raw and fitted ϕ–spectra plotted on a linear vertical scale. (c) and (d) show raw and fitted ϕ–spectra plotted on a dB scale.

9.5.2 Qualitative Discussion of the Results for the Two-tone Suppression Experiment

Fig. 9.15 shows measured ear canal pressure in both the partially suppressed and maximally suppressed conditions for the two-tone suppression experiments for all four subjects. (Results from only one session are shown.) This shows the expected pattern of ripples, with a spacing which corresponds to the SFOAE periodicity. The ripple amplitude reduces as the suppressor level is increased.

Fig. 9.16 shows various results derived from the measurement of the complex SFOAE pressure for subject 1, plotted against the logarithmic frequency variable, η_{Oct} for each probe level. The results appear to be qualitatively similar to those seen in the self-suppression experiment (fig. 9.7), except that now the maximum change in the normalised SFOAE level is less.

Fig. 9.17 shows the raw measured and the fitted ϕ-spectra obtained from the η-series shown in fig. 9.16. The fitted ϕ-spectra show a clear difference from those in the self-suppression experiment (fig. 9.8). Instead of showing a reduction in value, ϕ_C shows a tendency to increase in value as the normalised SFOAE amplitude reduces. The change is less marked than for the self-suppression experiment, but it is still clearly discernible. This change is in agreement with the predictions from the cochlear models based on Shera and Zweig's theory (fig. 7.5).

Figs. 9.18 and 9.19 show the SFOAE η-series and ϕ-spectra for subject 20. In fig. 9.19 the trend of increasing ϕ_C with reducing amplitude of the normalised SFOAE amplitude is seen even more clearly than for subject 1.

Figs. 9.20 and 9.21 show the SFOAE η-series and ϕ-spectra for subject 15. In fig. 9.21 the trend in ϕ_C less clear than for subjects 1 and 20.

The results for subject 3 are not shown, since all the frequency sweeps had a poor SNR and were therefore rejected.

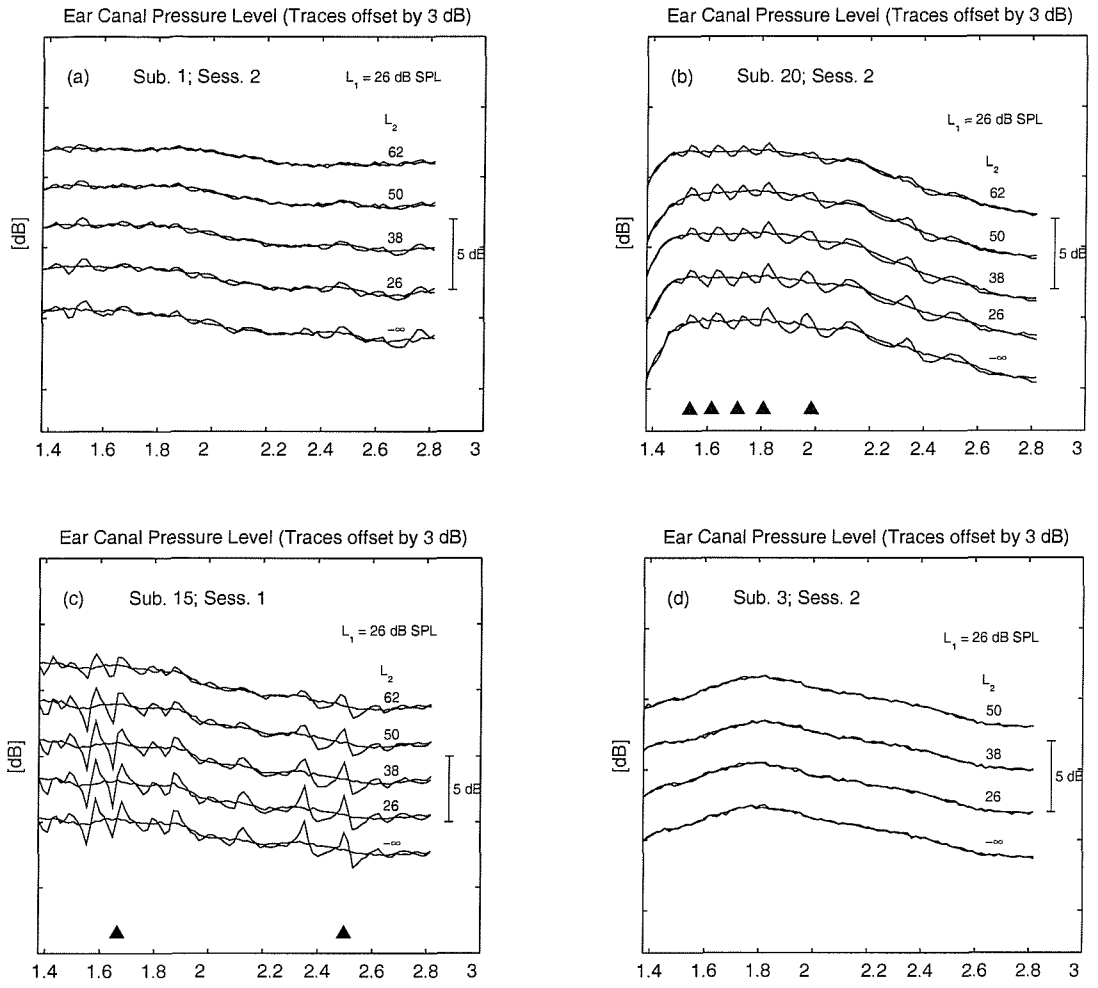


Fig 9.15: Ear canal sound pressure level for the two-tone suppression experiment for four subjects. The ear canal pressure at the stimulus frequency is shown. For clarity, traces are offset vertically, with the bar indicating the range. The results are from subjects with (a) moderate SFOAEs, but no detectable SOAE (subject 1; session 2); (b) strong SFOAE, and with both 'low' and 'high' S/EOAE strength ratings (subject 20, session 2); (c) strong SFOAE and with 'high' S/EOAE strength ratings throughout (subject 15; session 1); (d) weak SFOAE and with no detectable SOAE (subject 3, session 2). The partially suppressed ear canal pressure, p_{EC}, appears as a rippled trace, whilst the maximally suppressed ear canal pressure, $p_{EC:R=0}$, appears as a smooth trace. Solid triangles indicate any SOAE frequencies.

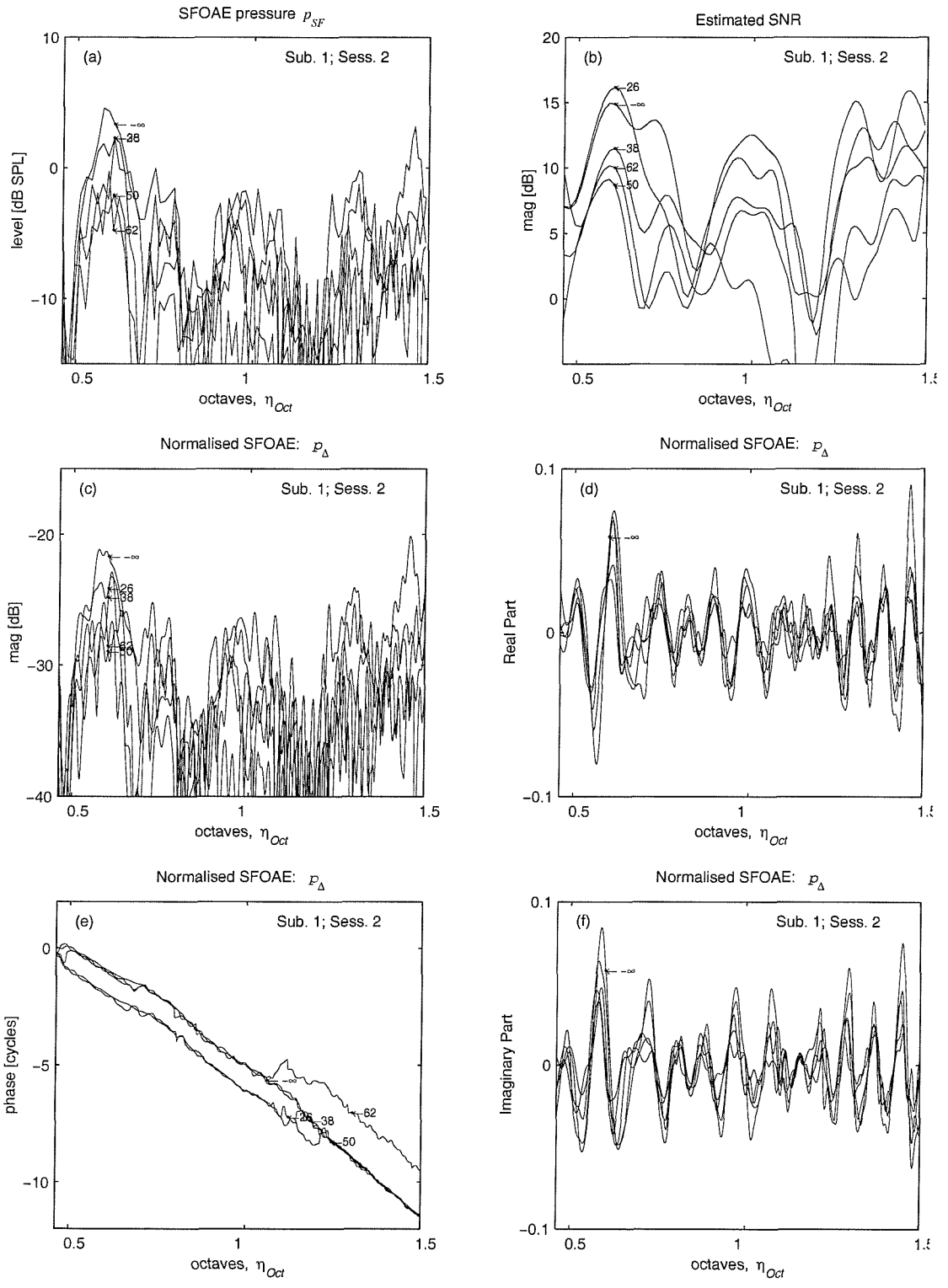


Fig 9.16 Measured SFOAE frequency sweep data for the two-tone suppression experiment, for a subject showing moderate SFOAEs, but with no detectable SOAEs (subject 1, session 2). The nominal probe tone level is held constant throughout at: L_1 =26 dB SPL. The nominal suppressor tone levels (indicated on curves) are: $L_2 = -\infty$, 26, 38, 50 and 62 dB SPL. (a) SFOAE pressure level; (b) estimated SNR; (c) normalised SFOAE level (d) real part of normalised SFOAE; (e) phase of normalised SFOAE; (f) imaginary part of normalised SFOAE.

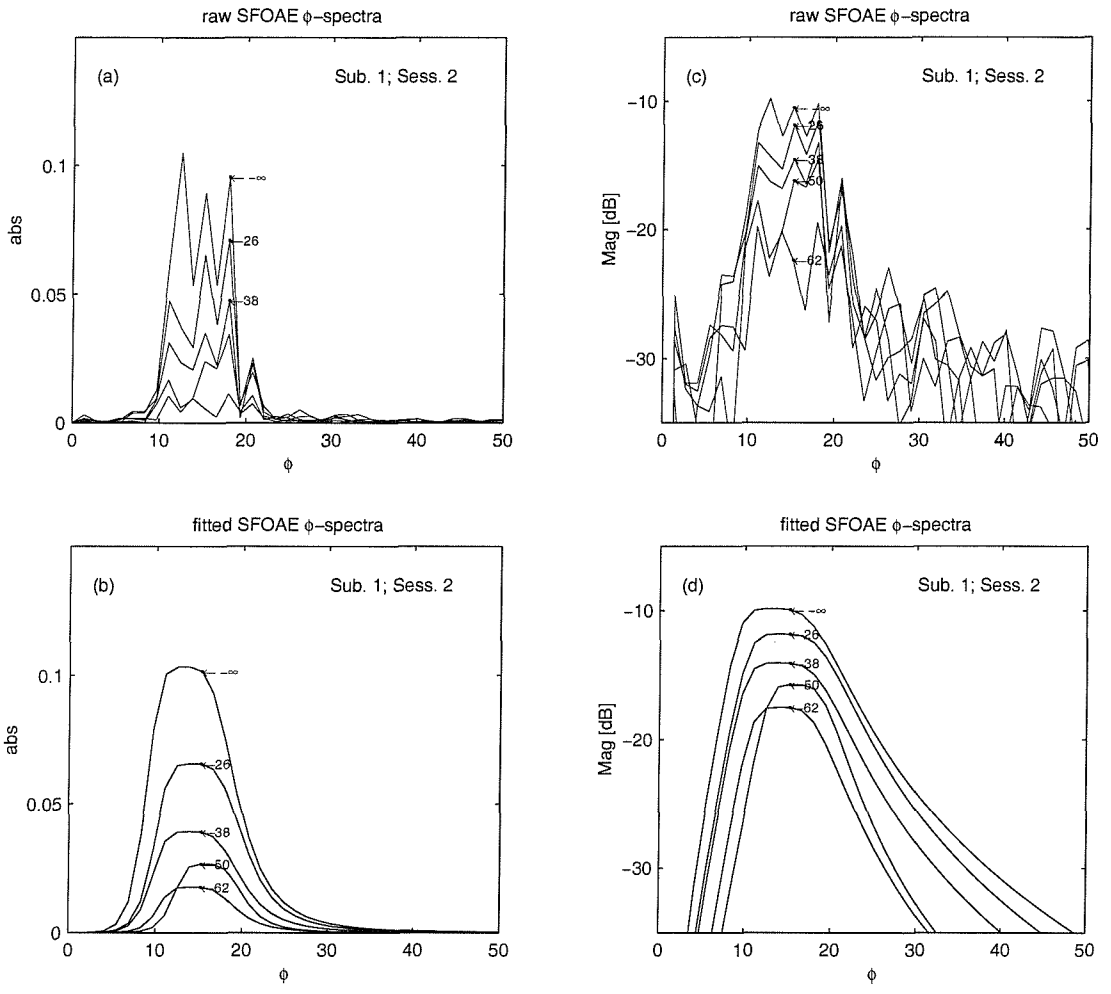


Fig 9.17 Raw and fitted SFOAE ϕ–spectra for the two-tone suppression experiment, for a subject showing moderate SFOAEs, but with no detectable SOAEs (subject 1, session 2). The nominal probe tone level is held constant throughout at: L_1 =26 dB SPL. The nominal suppressor tone levels (indicated on curves) are: $L_2 = -\infty$, 26, 38, 50 and 62 dB SPL.
(a) and (b) show raw and fitted ϕ–spectra plotted on a linear vertical scale. (c) and (d) show raw and fitted ϕ–spectra plotted on a dB scale.

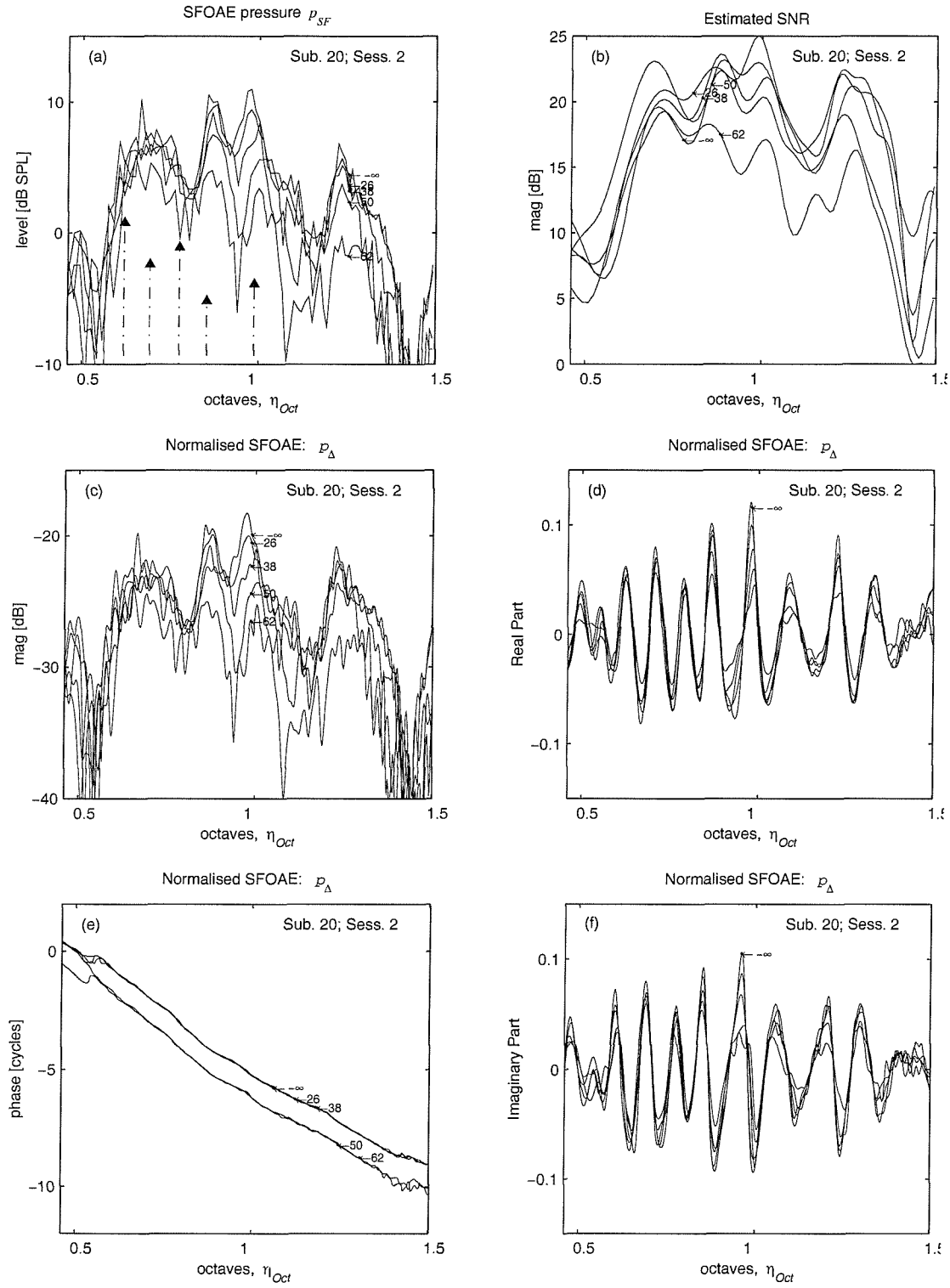


Fig 9.18 Measured SFOAE frequency sweep data for the two-tone suppression experiment, for a subject showing strong SFOAEs, and with both 'low' and 'high' S/EOAE strength ratings (subject 20, session 2). The nominal probe tone level is held constant throughout at: L_1 =26 dB SPL. The nominal suppressor tone levels (indicated on curves) are: $L_2 = -\infty$, 26, 38, 50 and 62 dB SPL. (a) SFOAE pressure level; (b) estimated SNR; (c) normalised SFOAE level (d) real part of normalised SFOAE; (e) phase of normalised SFOAE; (f) imaginary part of normalised SFOAE.

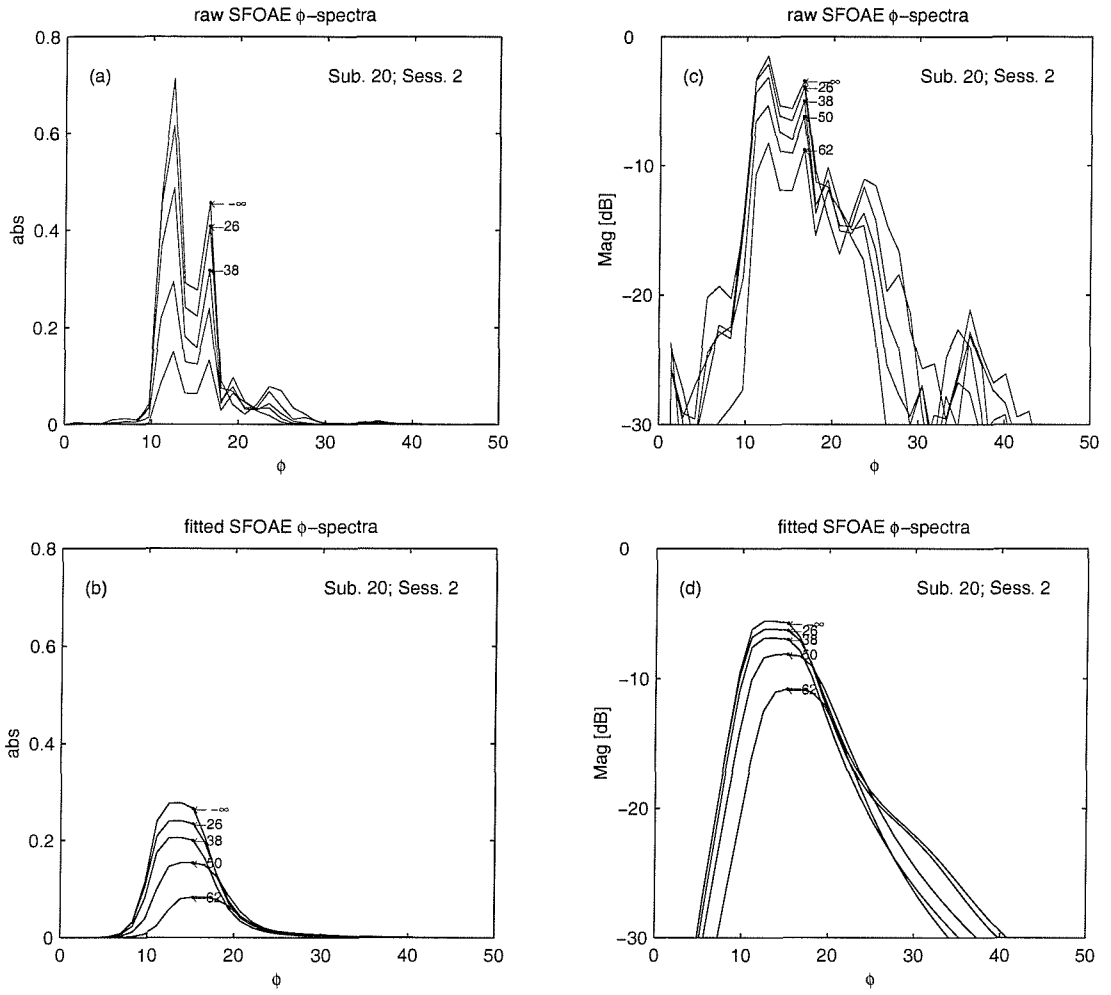


Fig 9.19 Raw and fitted SFOAE ϕ–spectra for the two-tone suppression experiment, for a subject showing strong SFOAEs, and with both 'low' and 'high' S/EOAE strength ratings (subject 20, session 2). The nominal probe tone level is held constant throughout at: L_1 =26 dB SPL. The nominal suppressor tone levels (indicated on curves) are: $L_2 = -\infty$, 26, 38, 50 and 62 dB SPL.
(a) and (b) show raw and fitted ϕ–spectra plotted on a linear vertical scale. (c) and (d) show raw and fitted ϕ–spectra plotted on a dB scale.

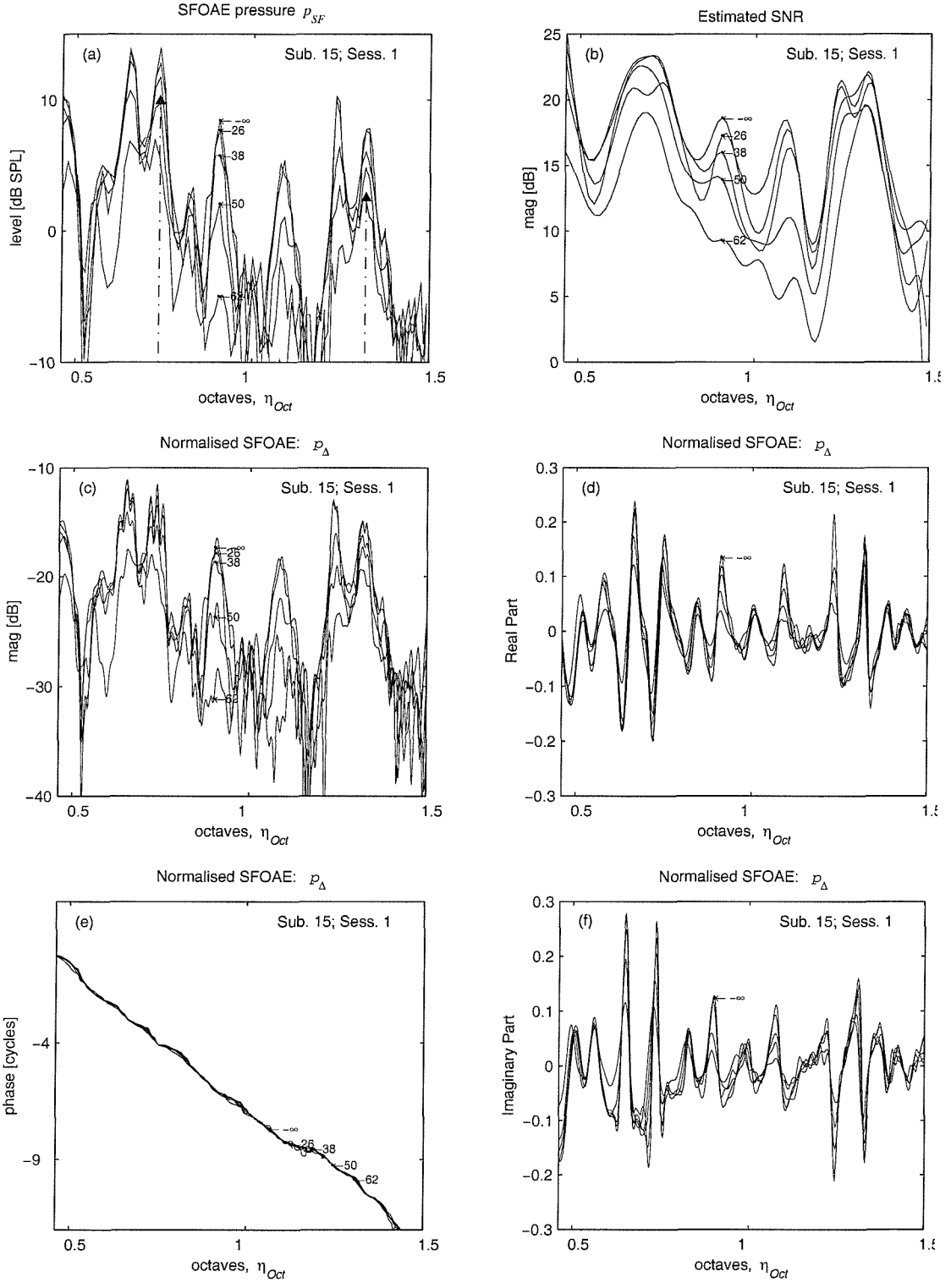


Fig 9.20 Measured SFOAE frequency sweep data for the two-tone suppression experiment, for a subject showing strong SFOAEs, and with 'high' S/EOAE strength ratings (subject 15, session 1). The nominal probe tone level is held constant throughout at: L_1 =26 dB SPL. The nominal suppressor tone levels (indicated on curves) are: $L_2 = -\infty$, 26, 38, 50 and 62 dB SPL. (a) SFOAE pressure level; (b) estimated SNR; (c) normalised SFOAE level (d) real part of normalised SFOAE; (e) phase of normalised SFOAE; (f) imaginary part of normalised SFOAE.

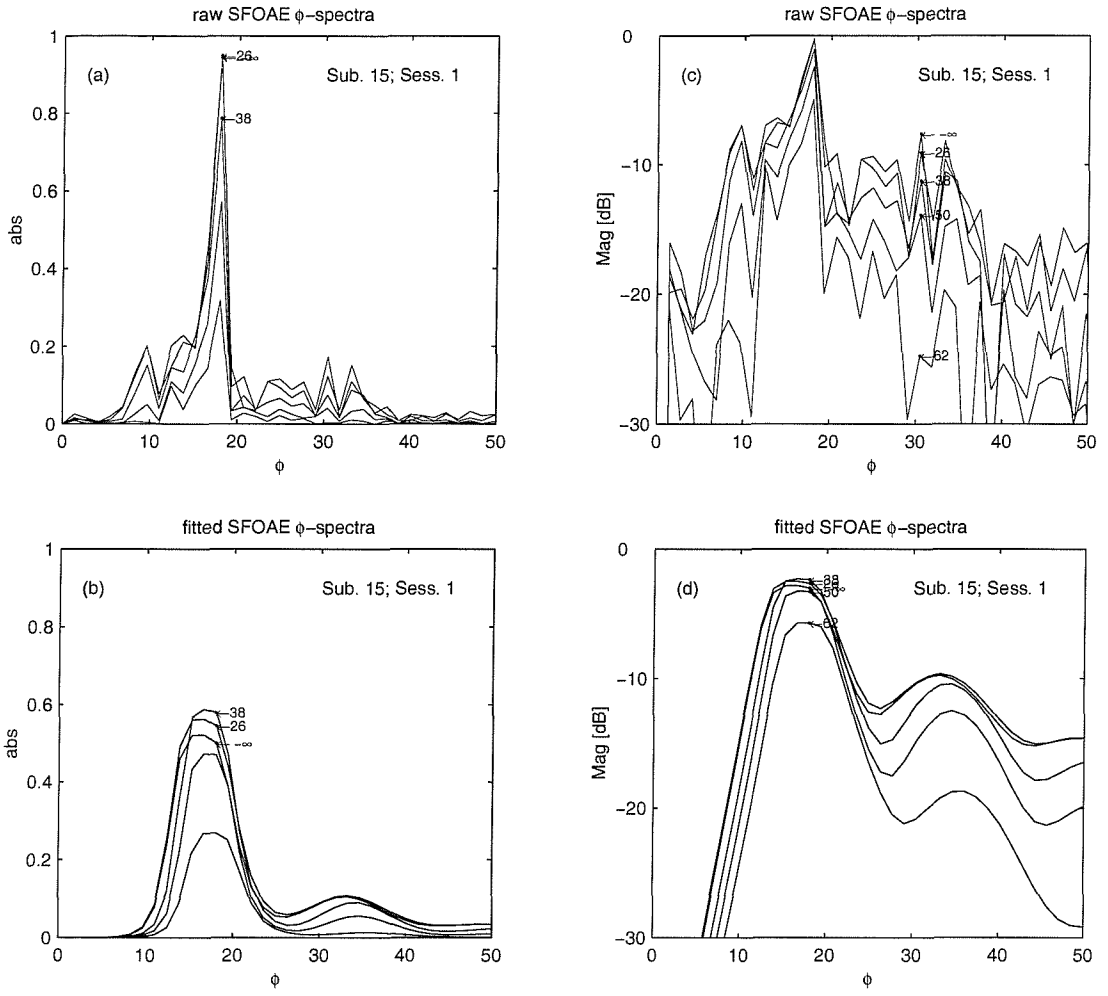


Fig 9.21 Raw and fitted SFOAE ϕ–spectra for the two-tone suppression experiment, for a subject showing strong SFOAEs, and with 'high' S/EOAE strength ratings (subject 15, session 1). The nominal probe tone level is held constant throughout at: L_1 =26 dB SPL. The nominal suppressor tone levels (indicated on curves) are: $L_2 = -\infty$, 26, 38, 50 and 62 dB SPL. (a) and (b) show raw and fitted ϕ–spectra plotted on a linear vertical scale. (c) and (d) show raw and fitted ϕ–spectra plotted on a dB scale.

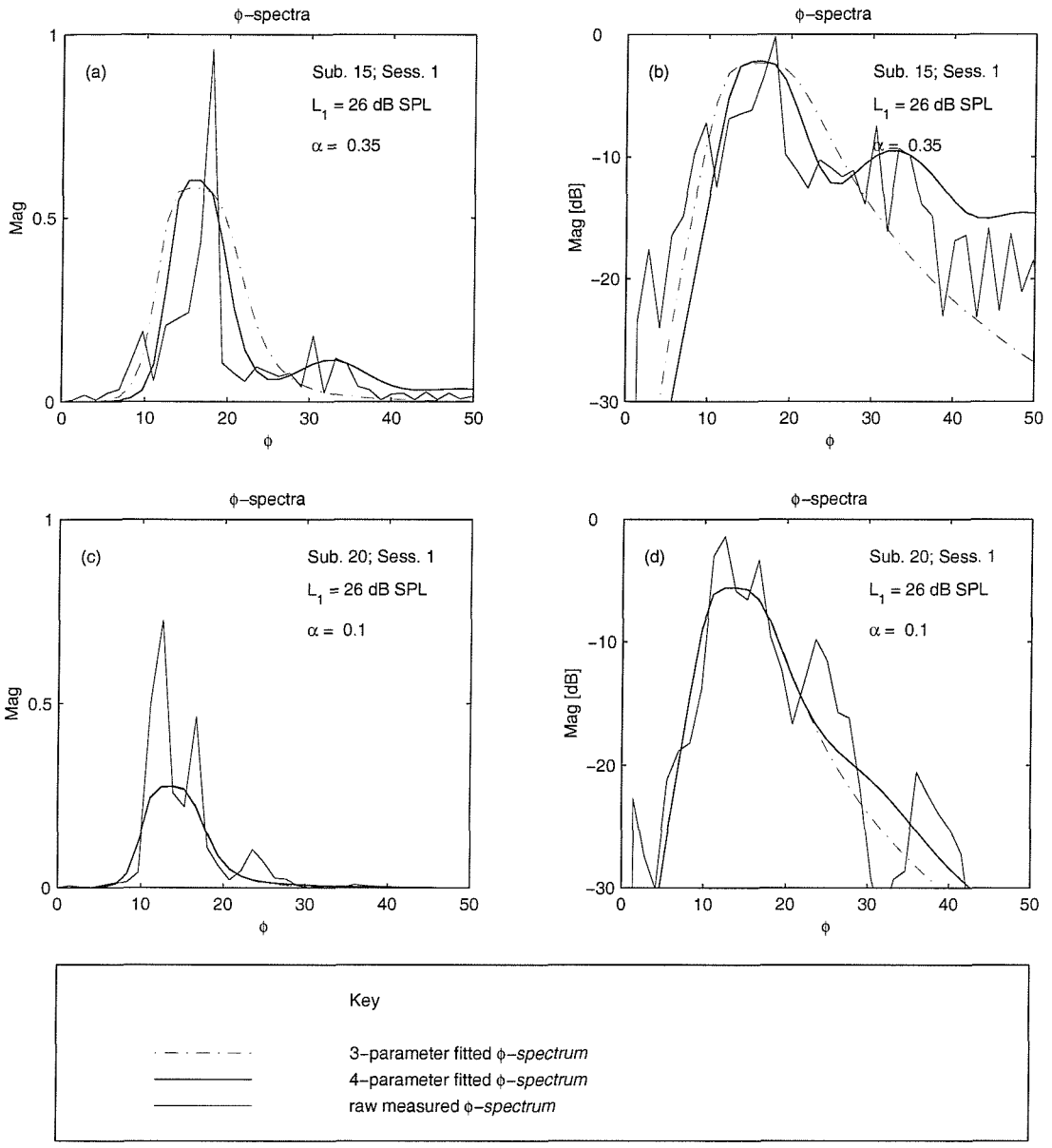


Fig 9.22 Comparison of the fitted ϕ–spectra arising from the 4-parameter model with that from the 3-parameter model. Results from two frequency sweeps are shown. (a) and (b) show the ϕ–spectra for a case where the 4-parameter model returns a value of $\alpha = 0.35$, which, if the model is correct, indicates significant multiple reflection (subject 15; session 1; self-suppression; $L_1 = 26$ dB SPL). The ϕ–spectra for the 3- and 4-parameter model show significant differences. (c) and (d) show the ϕ–spectra for a case where the 4-parameter model returns a value of $\alpha = 0.1$, thus predicting that multiple reflection are insignificant (subject 20; session 1; self-suppression; $L_1 = 26$ dB SPL). The ϕ–spectra for the 3- and 4-parameter model show no significant differences. Note that the fit is performed on the data in its linear, rather than decibel form, and therefore the representations in panels (a) and (c) give a more direct indication of the quality of the fit than panels (b) and (d).

9.5.3 Comparison of Results of the 3-Parameter and 4-Parameter Models.

Before looking at the detailed results from the 4-parameter model, the issue of the validity of the model is briefly examined. Questions about the validity of the model arose in the discussion of fig. 9.8, where the value of α did not appear to vary smoothly with probe level. Although the values of α are not the main focus of this investigation, it might be thought that errors in the estimates of α may be symptomatic of error in estimates of the other parameters.

In this section it is asserted that the estimates of α from the 4-parameter model are in general valid, though they suffer from a higher than desirable random error. This probably leads to an overestimate of α for some frequency sweeps. However, the effect of this random error on the accuracy of estimates of ϕ_C is expected to be very slight. The effect on estimates of ϕ_{BW} is expected to be greater, probably leading overestimates in ϕ_{BW} for certain frequency sweeps.

Evidence for this assertion comes from the application of both the 4-parameter model and 3-parameter model (which is equivalent to the 4-parameter model with α set to zero) to the results, which led to the following observations being made.

- For small values of α ($\alpha < 0.15$) the effect of α on the other three parameters: ϕ_C, ϕ_{BW} and β is negligible. This means that virtually the same results are obtained with the 3-parameter and 4-parameter models.
- For larger values of α ($\alpha > 0.15$) the effect of α on ϕ_C and β is only slight, whilst the effect on ϕ_{BW} is more significant.

These two points are illustrated in fig. 9.22, which compares the results for the 3-parameter and 4-parameter models. Two cases are examined, the first where a high value of α is returned by the 4-parameter model; the second where a low value of α is returned. For the high value of α (panels a and b), the fitted ϕ-spectra for the 3- and 4-parameter models are clearly different (though the effect on ϕ_C is only slight). For the low value of α the fitted ϕ-spectra for the 3- and 4-parameter models are indistinguishable when plotted on a linear scale (panel (c)). Some difference is seen when plotted on a dB scale, but there is no significant difference in the estimated values of ϕ_C and β. From results such as these it was concluded that ϕ_C is largely insensitive to errors in the estimate of α.

Evidence for the validity of the estimates of α comes from applying the 4-parameter model to all the frequency sweeps in a level series. This showed that, when the value of α exceeds about 0.15, its variation within the level series is, on average, in accordance with

predictions from the cochlear models. However, the value of α also shows significant, apparently random, scatter around this central trend.

This result is illustrated and discussed in greater detail in section 9.6.6. The result lends weight to the argument that α gives a useful measure of the physical phenomenon of multiple reflection in the cochlea. The error in estimates of α, which is suggested by the erratic trends across some of the level series, is thought to be a consequence of the random process underlying SFOAE generation. Support for this also comes from the study of cochlear models in section 6.8, where the performance tests on the 4-parameter model using the cochlear models showed that the estimates of α and ϕ_{BW} suffered from significant random error.

The conclusion from these results is that the 4-parameter model gives more reliable estimates than the 3-parameter model, though significant random errors in the estimates of α on ϕ_{BW} expected. Errors ϕ_C arising from error is α are expected to be slight.

9.6 Variations across the Level Series

In this section, variations across the level series are examined. Thus for each frequency sweep, a single set of characterising quantities are calculated, most important of which are the parameters returned by the 4-parameter model. The variation of these quantities can then be traced for both the self-suppression and the two-tone suppression experiments.

9.6.1 I/O Functions: Variation of SFOAE Level

For each frequency sweep, the strength of the SFOAE pressure has been characterised by calculating an RMS value of the SFOAE pressure, p_{SF}, across the frequency range. This is the quantity denoted by S_1 is equation [9.8], and will be referred to as the RMS SFOAE pressure, or, when expressed in dB, the RMS SFOAE pressure level. The variation of this value across the level series for both the self-suppression and the two-tone suppression experiments is shown in fig. 9.23 for each of the 20 subjects. The variations are shown against either the nominal probe level or the nominal suppressor level, as appropriate. Both sessions for each experiment are shown. The figure also indicates which frequency sweeps have been rejected, based on the SNR_1 estimate. Also indicated are the S/EOAE ratings.

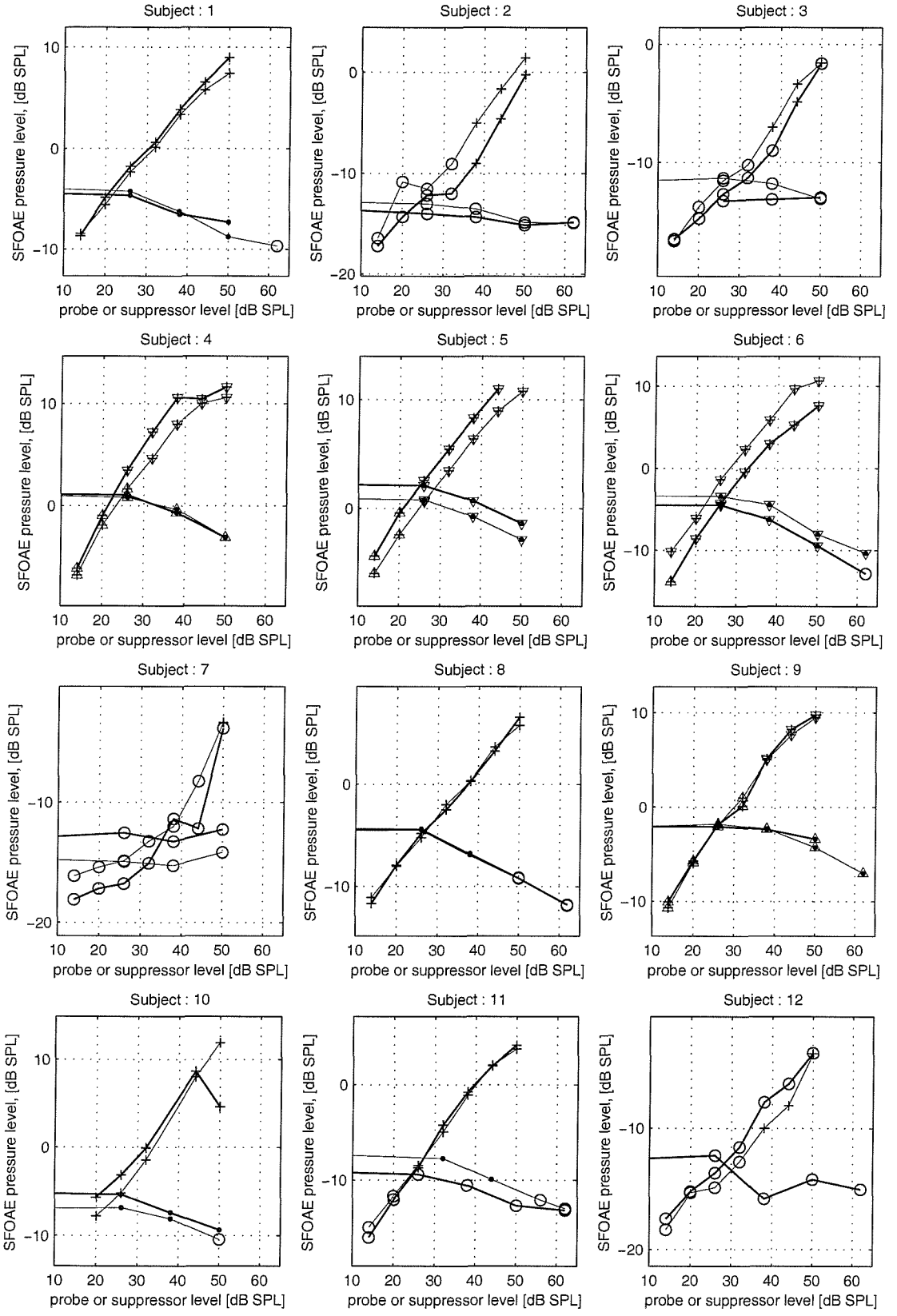


Fig 9.23 Variation of the measured rms SFOAE pressure level during both the self-suppression and the two-tone suppression experiments. Each point represents the $p_{SF:RMS}$ value obtained by frequency averaging $|p_{SF}(f)|^2$ over one sweep. The horizontal axis shows the nominal level of the probe tone (self-suppression experiment) or the suppressor tone (two-tone suppression experiment). Results are shown for all 20 subjects and for both sessions. The S/EOAE rating of each sweep is also indicated. Sweeps that will be rejected due to poor SNR are also shown. Key shown on the following page. Continued over page.

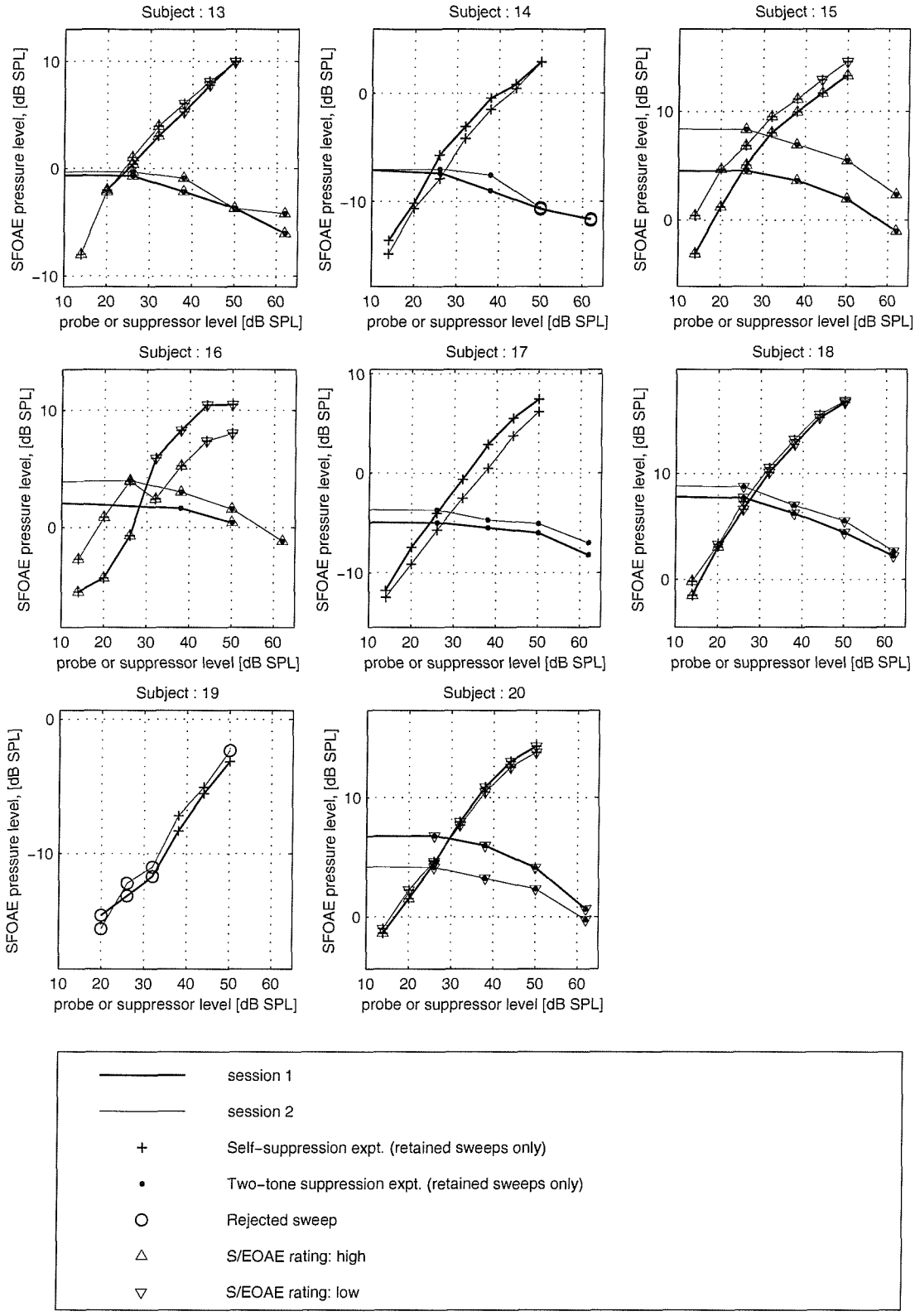


Fig 9.23 continued.

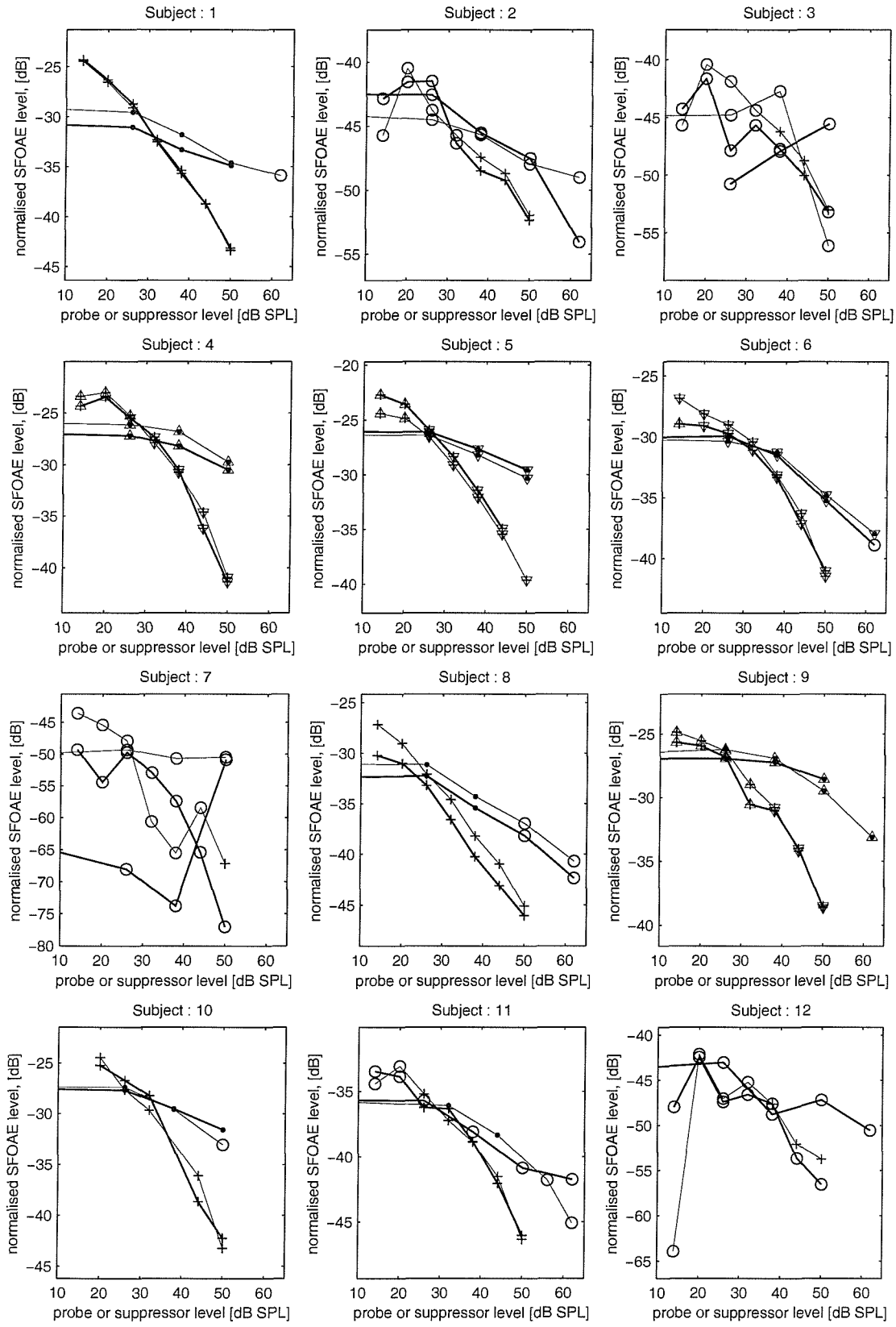


Fig 9.24 Variation of the measured RMS normalised SFOAE level during both the self-suppression and the two-tone suppression experiments. Each point represents the $p_{\Delta:\mathrm{RMS}}$ value obtained by frequency averaging over one sweep. The horizontal axis shows the nominal level of the probe tone (self-suppression experiment) or the suppressor tone (two-tone suppression experiment). Results are shown for all 20 subjects and for both sessions. The S/EOAE rating of each sweep is also indicated. Sweeps that will be rejected due to a poor SNR are also shown. Key shown on the following page. Continued over page.

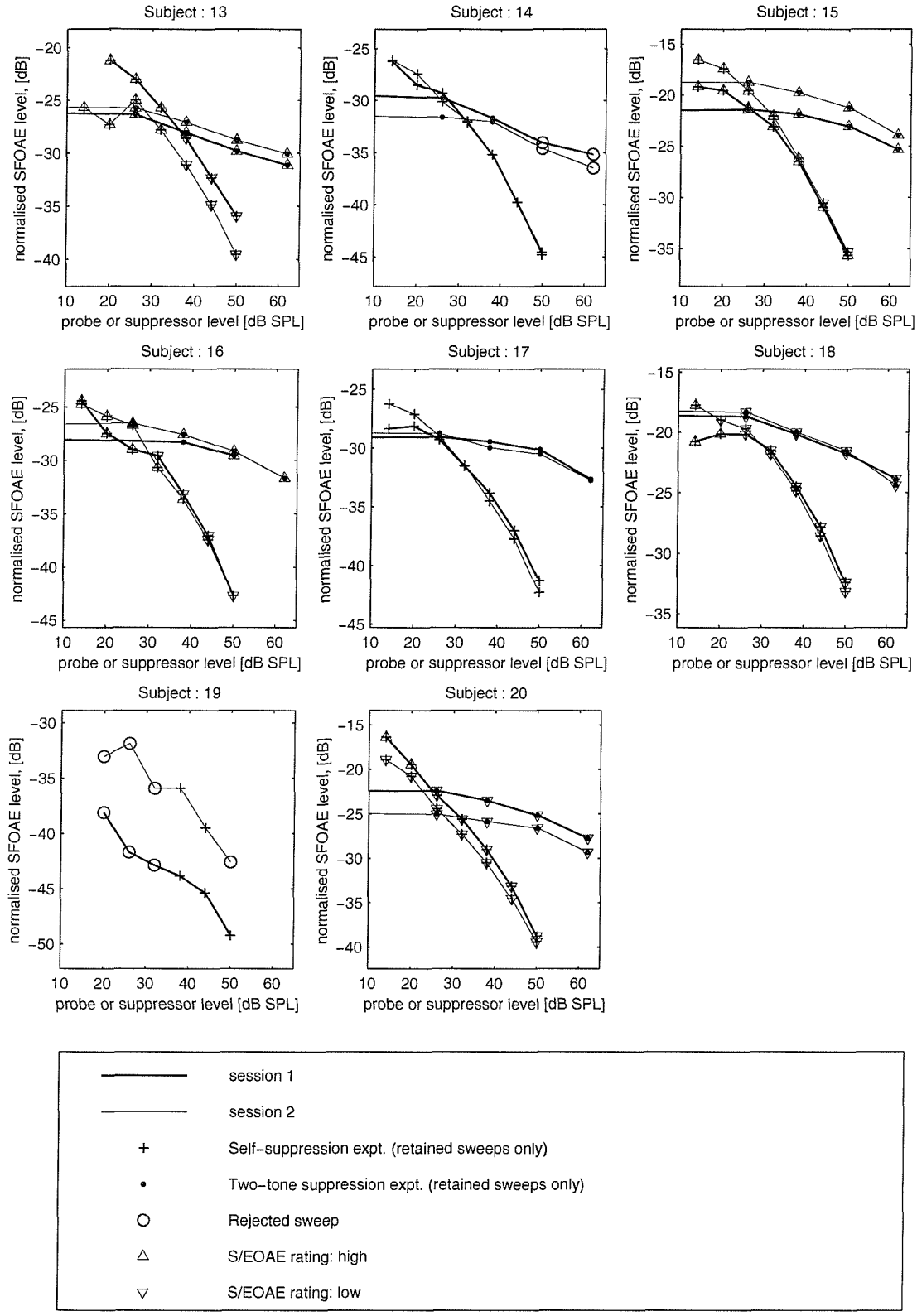


Fig 9.24 continued.

The results show in general that the RMS SFOAE pressure level grows at approx. 0.5 dB per dB increase in probe level. Maximum levels of around 12 dB SPL are obtained. Frequency sweeps with a level below –10 dB SPL usually become rejected on the basis of their poor SNR_1 estimate. Note that it is not safe to infer the presence of a signal simply from an increase in SFOAE with probe level, because the estimate of the SFOAE level will contain a component due to noise. Because of the reduction in the averaging time with probe level, the noise level will also tend to increase with probe level. The variation of noise is discussed in the next section.

Typically, at the maximum suppressor tone levels, the RMS SFOAE pressure level is reduced by about 7 dB, relative to its unsuppressed value.

The variation of the RMS normalised SFOAE level, $L_{\Delta:RMS}$, for each frequency sweep is presented in fig. 9.24. This is calculated from the RMS average of the normalised pressure, p_Δ, across logarithmic frequency giving $p_{\Delta:RMS}$ (equation [6.10]). In fact this value is virtually identical to the parameter β, returned by the 4-parameter model (section 6.6). These curves show the deviation from nonlinearity, since, for a linear system, the normalised SFOAE would remain unchanged with changes in probe or suppressor level. The maximum values of the frequency averaged normalised SFOAE are around –15 dB. Typically the values fall at about 0.5 dB per dB increase in probe level.

9.6.2 Variation of SNR

In order to show clearly which frequency sweeps have been rejected for each subject, the variation of the SNR_1 estimate over the level series is shown in fig. 9.25. Instead of plotting the data against probe or suppressor level, the variation in SNR_1 is plotted against the RMS normalised SFOAE level, $L_{\Delta:RMS}$, which, for brevity, will be referred to as the 'normalised SFOAE level'. The reason for adopting this value as the abscissa is that it allows both the self-suppression and the two-tone suppression experiments to be plotted against a common axis (cf. figs. 7.6 and 7.7). This turns out to be particularly useful for future plots, and for comparison with model results. The plot shows where SNR_1 values fall below 6 dB, leading to the sweep being rejected.

For the self-suppression experiment, there is a tendency for SNR_1 to reduce as the normalised SFOAE level increases. This is because a high normalised SFOAE level corresponds to a low SFOAE pressure level, where it becomes uneconomic in terms of experimental time to attempt to average to the same SNR as at high SFOAE levels (section 8.8.3).

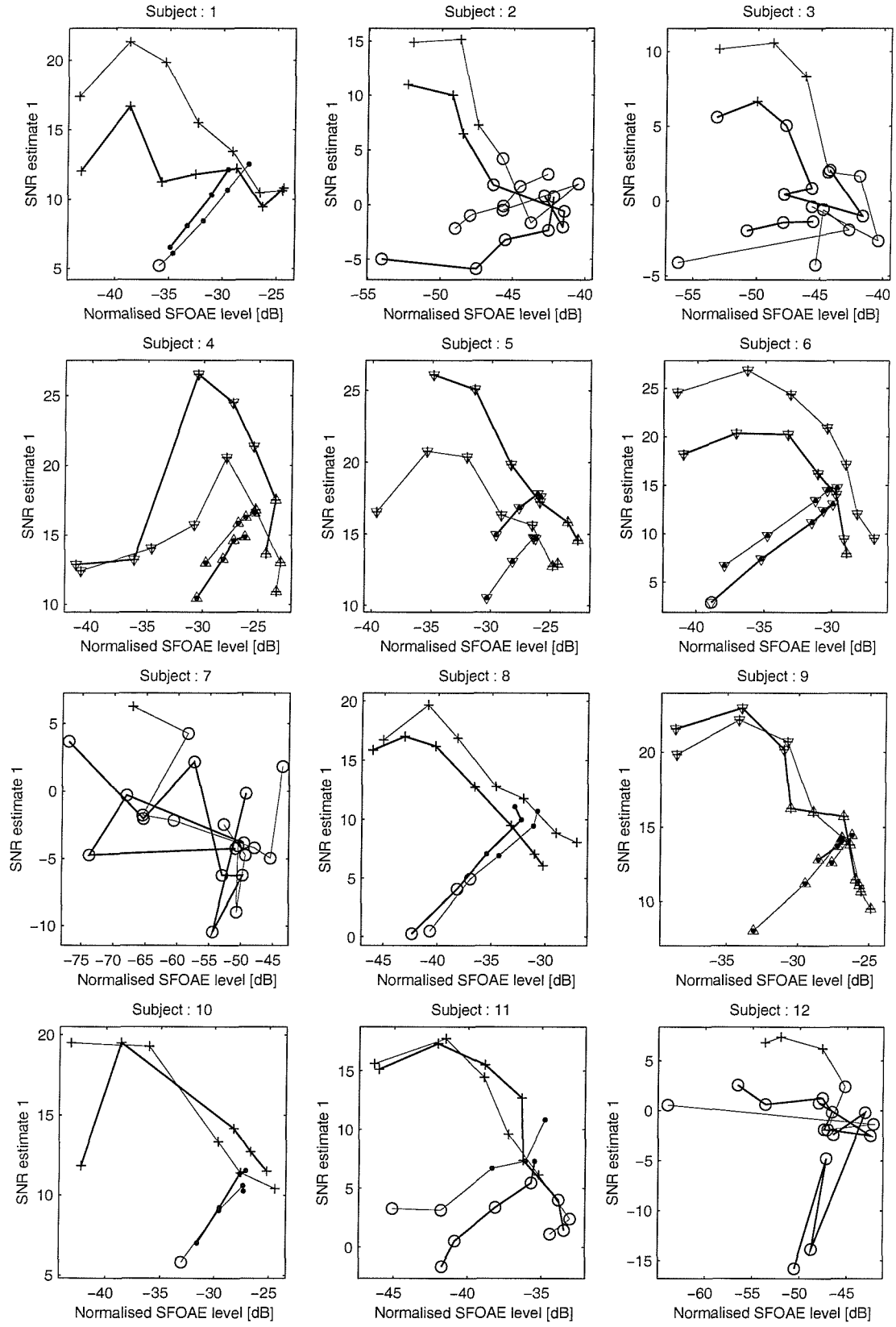


Fig 9.25 Variation of the estimated signal-to-noise ratio (the SNR_1 estimate) during both the self-suppression and the two-tone suppression experiments. Each point represents SNR_1 obtained from one sweep. The horizontal axis shows the normalised SFOAE level, $p_{\Delta:RMS}$. Results are shown for all 20 subjects and for both sessions. The S/EOAE rating of each sweep is also indicated. Key shown on the following page.

Continued over page.

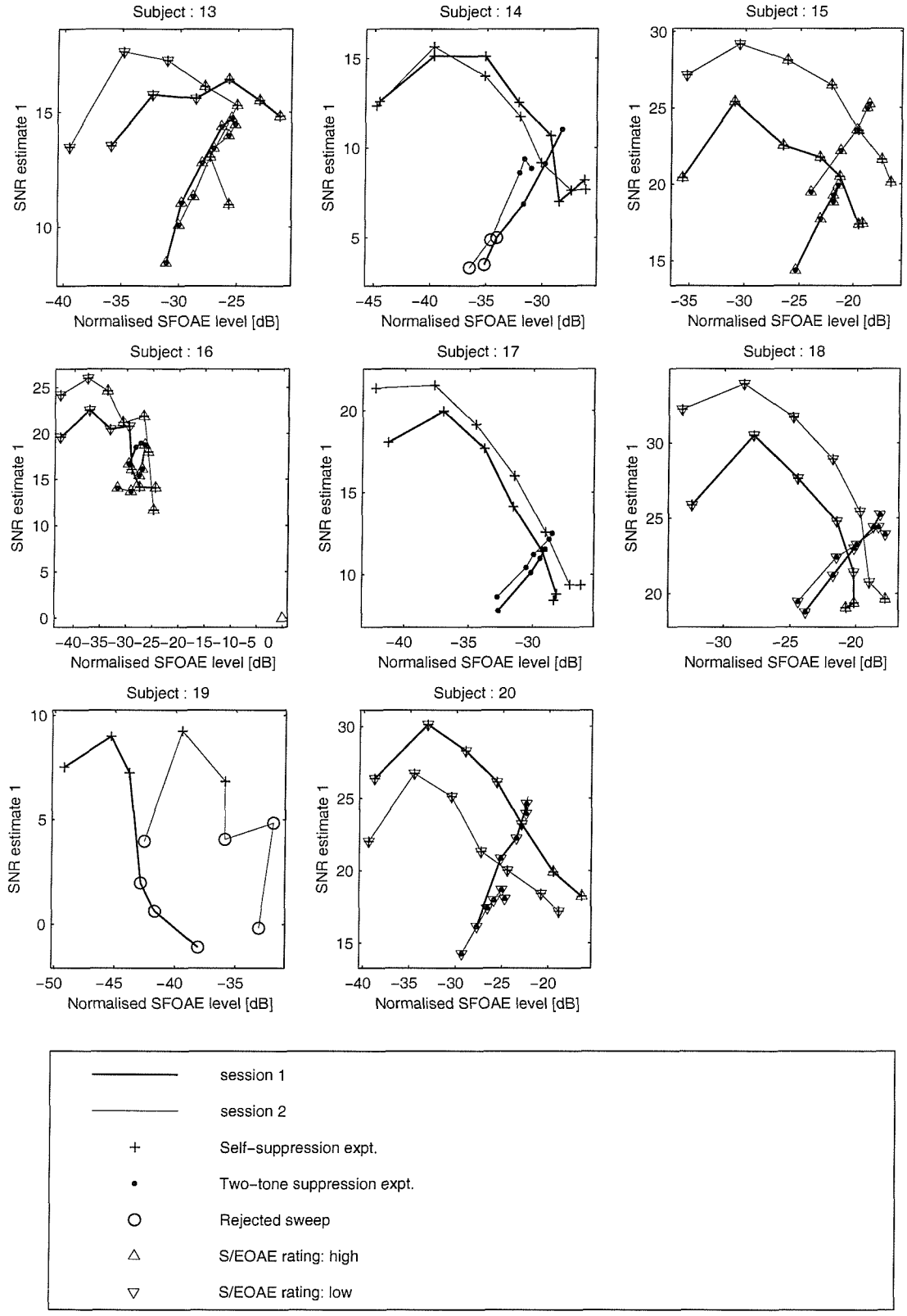


Fig 9.25 continued.

For the two-tone suppression experiment, SNR_1 increases with normalised SFOAE level. This is because all results were averaged to the same estimated noise level, irrespective of the suppressor level. Therefore SNR_1 is reduces as the SFOAE pressure level (and therefore also the normalised SFOAE level) is reduced by an increase in suppressor level.

9.6.3 Variation of ϕ-centre value

Fig. 9.26 shows the variation in the ϕ-centre value, ϕ_C , with the normalised SFOAE level for all 20 subjects. The results from frequency sweeps that were rejected on the basis of their poor SNR_1 value are not shown. The S/EOAE rating is shown for each frequency sweep. It is useful to bear in mind the corresponding curves from the cochlear models, shown in figs. 7.6. and 7.7. These curves resemble the character "y", where the longer limb of the "y" (which has a positive gradient) arises from the self-suppression experiment, and the shorter limb (which has a negative gradient) arises from the two-tone suppression experiment.

This predicted y-shape is seen clearly in some of the data (e.g., subject 20). In other cases the trends are less clear (e.g., subject 11), or there are too few acceptable frequency sweeps for any trends to be assessed (e.g., subject 3). It is also desirable to take account of the S/EOAE rating for the results.

In order to make a systematic comparison between the measured data for each experiment and the model predictions laid out in section 7.4, the curves from sessions 1 and 2 have been averaged to produce a single joint curve. Prior to averaging, any ϕ_C estimates that are deemed to be unreliable are rejected. This judgement of reliability is made based on the repeatability of the estimates between sessions 1 and 2. However, since these estimates were made on different days, with different probe fits, there may be significantly different excitation levels in the cochlea for the two sessions. Therefore, it is inappropriate to compare the absolute values of estimates across sessions. Instead, the trends for sessions 1 and 2 are compared. The precise procedure is as follows.

1. For each experiment and for each subject, the curve in fig. 9.26 for session 1 is compared with that for session 2. Any portions of the two curves which show different trends in both sessions are identified. For example, for subject 16 the portions of the curves for the self-suppression experiment defined by the three points obtained with L_1 = 14, 20 and 26 dB SPL have a very different trend in the three ϕ_C estimates for

session 1 compared to session 2. Such estimates (in both the sessions) are then classed as "unrepeatable" and are rejected from this analysis. All the remaining estimates are classed as "repeatable" and are included in the further analysis. Thus, in this example for subject 16, the estimates for L_1 = 30, 36, 42 and 50 dB SPL are retained. Note that the classification is based on the trends within the session, and not on the absolute difference in the estimate across sessions. Also, differences in trend that are judged to be quite small are not used to reject estimates. For example, for subject 4, the estimates for session 1, at L_1 = 30, 36 and 42 dB SPL do not define a monotonic curve, whilst the corresponding estimates for session 2 do. However in this case, it is judged that the difference is small, and hence all estimates are accepted for further analysis. Although in these cases there is clearly a certain degree of subjective assessment required, they are rare, and have not proved problematic.

2. In cases where an estimate is present for only one of the sessions, such as for subject 5, session 2, L_1 = 50 dB SPL, this estimate is rejected.
3. The number of repeatable estimates in each group is counted to establish whether there are sufficient data for further assessment.
4. The trend of ϕ_C with normalised SFOAE level in each group is classified, according to the following scheme:

   "As model": the measured trend follows the model predictions (section 7.4)

   "Anomalous:" the trend is not as predicted in section 7.4

   "Insufficient data": too few (<3) reliable estimates in curve to assess the trend.

The averaged curves are shown in fig. 9.27. Note that now two stages of data rejection have been applied. In the first, frequency sweeps were rejected based on their SNR_1 estimates. The 4-parameter estimates for all the retained frequency sweeps were then calculated. The marked points in fig. 9.27 correspond to the estimates for ϕ_C. The second stage of data rejection applies only to the averaged curve shown in the figure. In this stage, ϕ_C estimates are rejected, based on the variability between session 1 and session 2 using the procedure described above. One cause of this variability was probably subject movement, leading to degradation of the probe fit. A second possible cause is drift noise, which is not accounted for in the SNR_1 estimate.

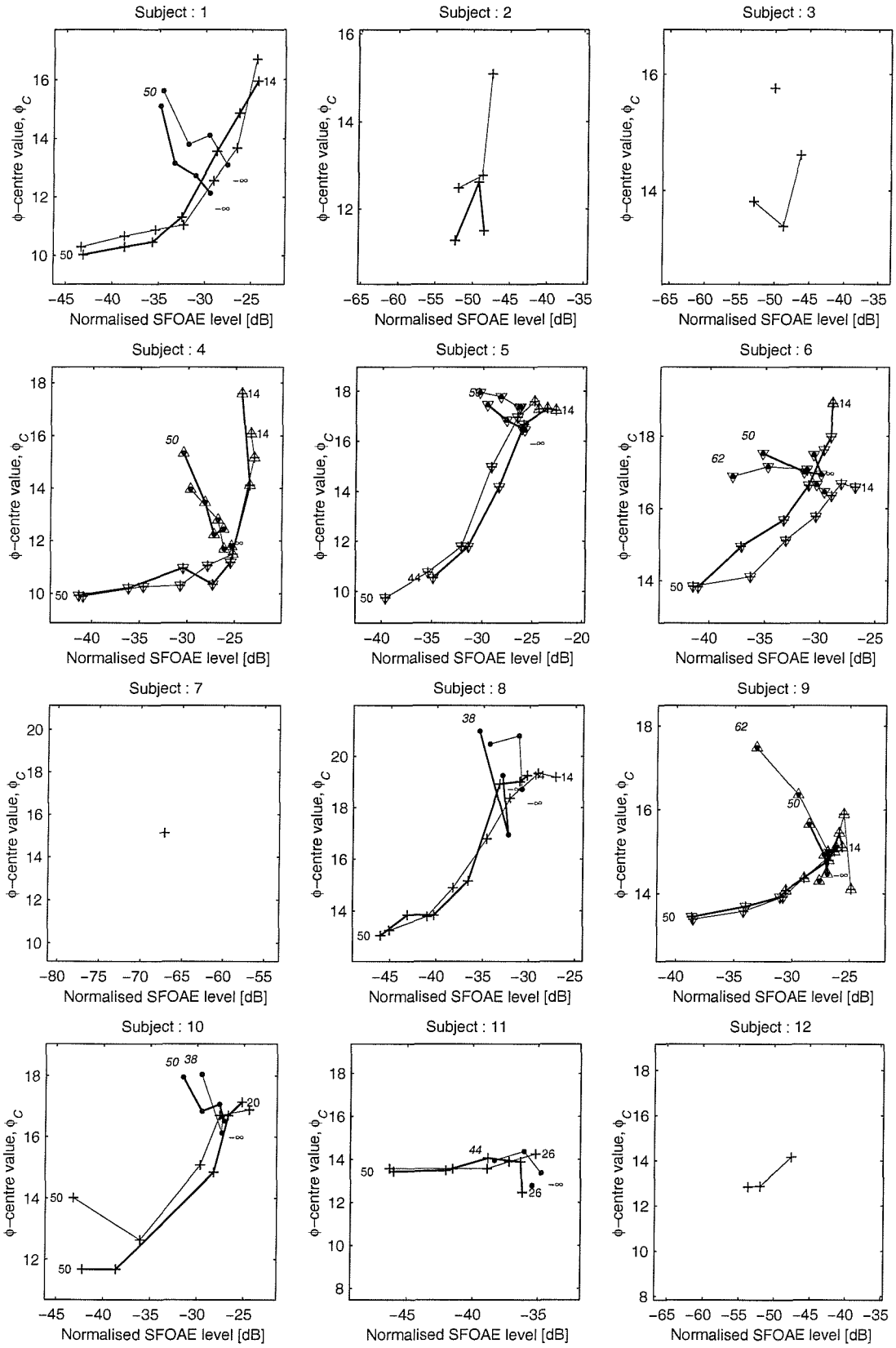


Fig 9.26. Variation of the measured SFOAE ϕ–centre value during both the self-suppression and the two-tone suppression experiments. Each point represents ϕ_C obtained from one sweep. The horizontal axis shows the normalised SFOAE level, $p_{\Delta:\mathrm{RMS}}$. Results are shown for all 20 subjects and for both sessions. The S/EOAE rating of each sweep is also indicated. Rejected sweeps are not shown. Numbers in normal font next to selected points indicate the nominal probe level for the self-suppression experiment, for that point. Numbers in italics next to selected points indicate the suppressor level for the two-tone suppression experiment, for that point. Key shown on the following page. Continued over page.

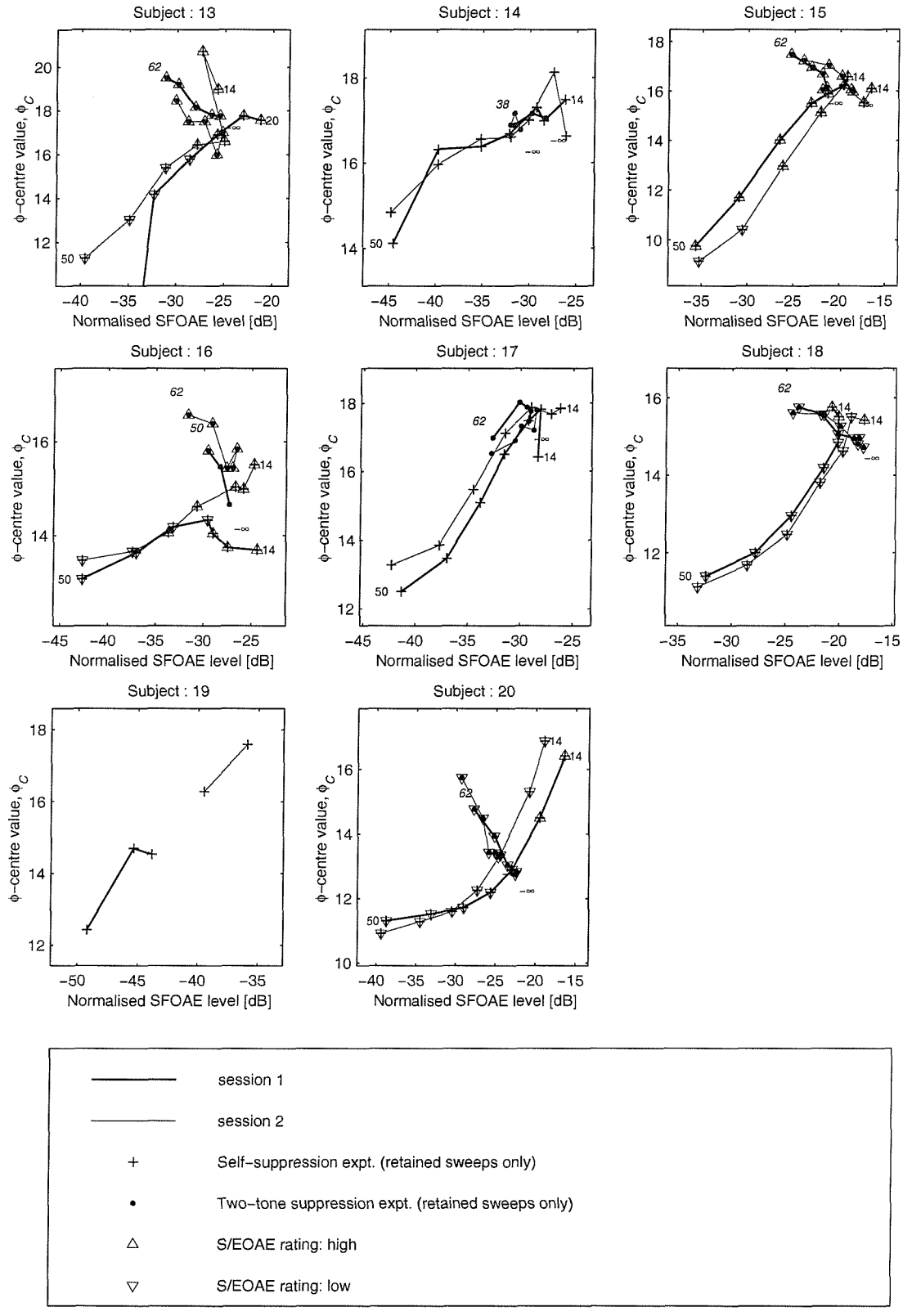


Fig 9.26 continued.

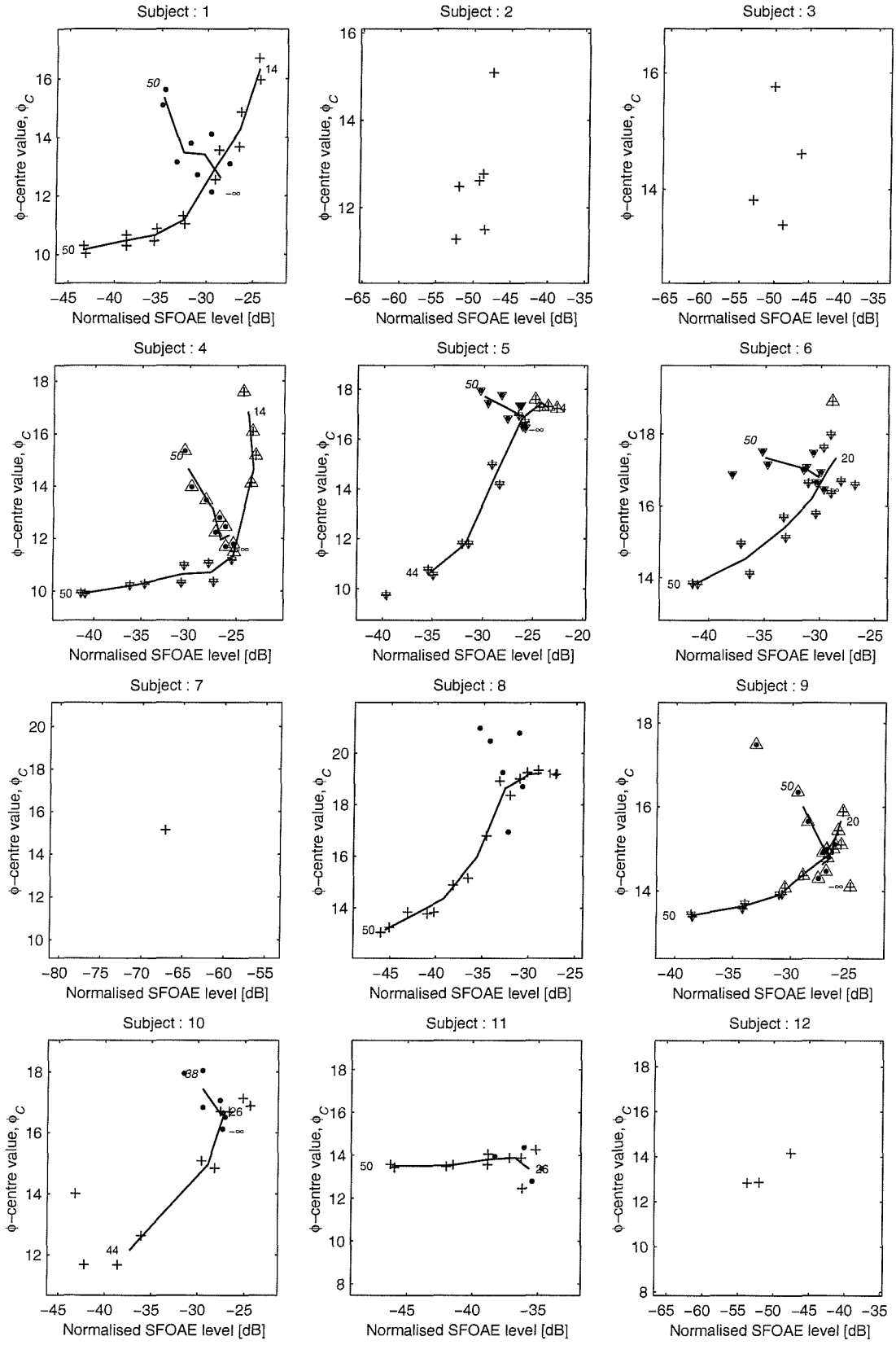


Fig 9.27 Measured ϕ–centre value curves defined by the average of estimates for session 1 and session 2. The variation is shown for both the self-suppression and the two-tone suppression experiments. The marked points are the actual estimates of ϕ_C for each sweep. The thick line passes though the average of the two estimates for session 1 and 2. Only accepted averages have been used. See text for the acceptance criteria. The horizontal axis shows the normalised SFOAE level, $p_{\Delta:\mathrm{RMS}}$. The S/EOAE rating of each sweep is also indicated. Rejected sweeps are not shown. Numbers in normal font next to the start and end of a curve indicate the nominal probe levels at the end points for the self-suppression experiment. Numbers in italics next to the start and end of a curve indicate the nominal suppressor level at the end points for the two-tone suppression experiment. Key shown on the following page. Continued over page.

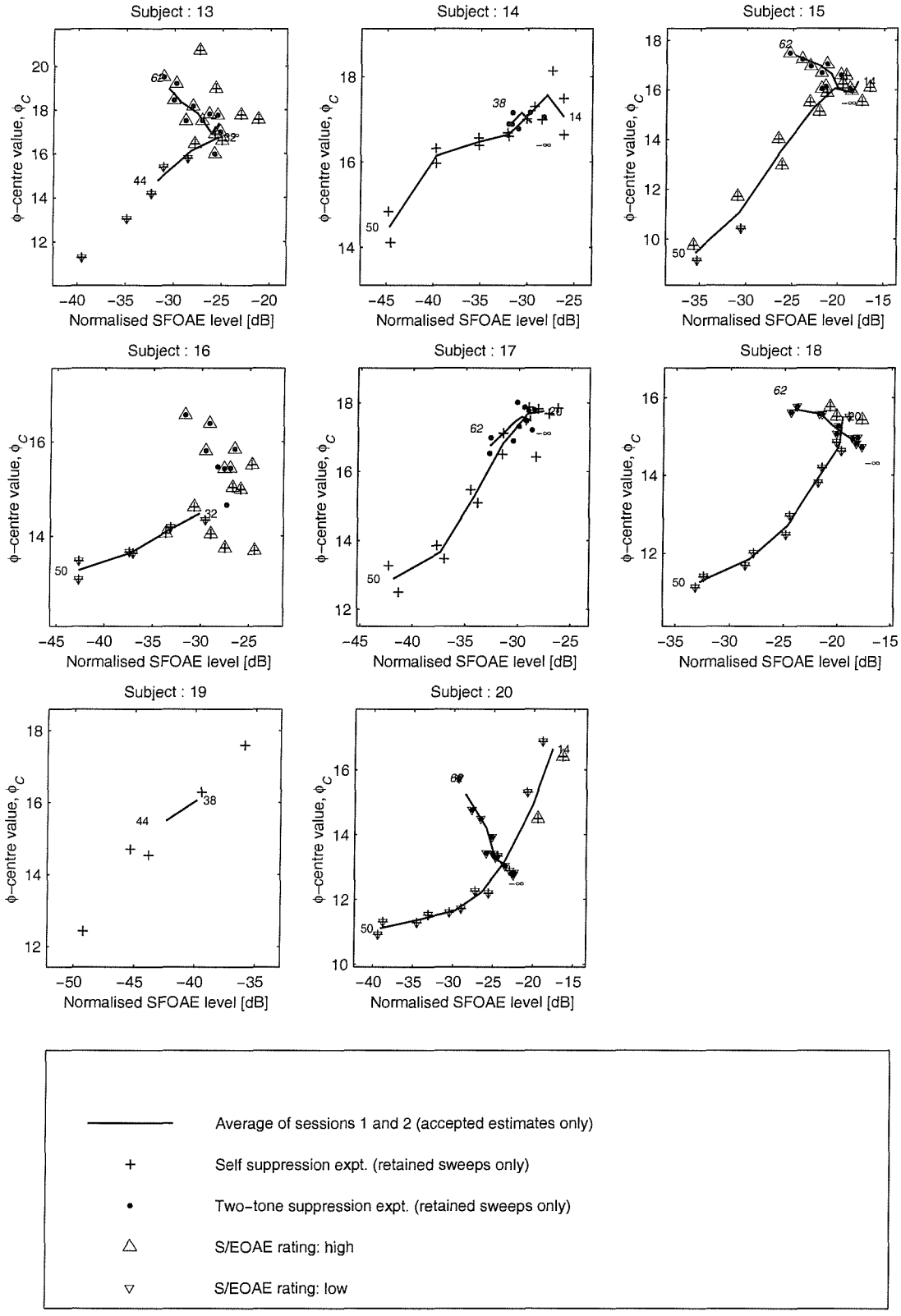


Fig 9.27. continued.

This summary results are shown in table 9.3:

Table 9.3: Summary of results for the measured ϕ_C variation with SFOAE level

Experiment	Classification of Trend	Number of Subjects	Subject Identification Numbers
Self-suppression	As model	14	1, 4-6, 8-10, 13-18, 20
	Anomalous	1	11
	Insufficient data	5	2, 3, 7, 12, 19
Two-tone suppression	As model	10	1, 4-6, 9, 10, 13, 15, 18, 20
	Anomalous	2	14, 17
	Insufficient data	8	2, 3, 7, 8, 11, 12, 16, 19
Both experiments	As model	10	1, 4-6, 9, 10, 13, 15, 18, 20

For the self-suppression experiment, only one result was classed as anomalous (subject 11). For this subject, ϕ_C varies little with SFOAE level. However, only the tail end of the curve was measurable (i.e., at high probe levels), where the predicted curve becomes quite flat. Therefore, this result is not strong evidence against the model.

For the two-tone suppression experiment, two results were classed as anomalous. (subjects 14 and 17). However, for subject 14, the reliable portion of the curve is very short, showing a change in normalised SFOAE level of only 2 dB. The results for subject 17 are more surprising, showing as a downward trend in ϕ_C over a 4 dB change in normalised SFOAE level.

The presence of SOAEs appears to have little influence on the trends. For the self-suppression experiment, results classed as "as model" were obtained where:

1. no SOAEs were detected (subjects 1, 8, 10, 14 and 17),
2. the S/EOAE rating was low for most sweeps (subjects 5, 6, 13, 16, 18, and 20)
3. the S/EOAE rating was high for most sweeps (subjects 4, 9, and 15).

Similarly for the two-tone suppression experiment, results classed as "as model" were obtained where:

1. no SOAEs were detected (subjects 1 and 10),
2. the S/EOAE rating was low for most frequency sweeps (subjects 5, 6, 13, 18, and 20),
3. the S/EOAE rating was high for most frequency sweeps (subjects 4, 9, and 15).

Within the "as model" class, a variety of different y-shapes are seen. It is speculated that subjects within this class show a variety of different cochlear amplifier characteristics,

leading to significant differences in the way in which the TW shape is altered during acoustic suppression.

In conclusion, the results give strong support to the model predictions made for the self-suppression experiment. The results for the two-tone suppression are less clear cut, but are, in general, in agreement with model predictions.

9.6.4 Variation of ϕ-bandwidth

Fig. 9.28 shows the variation in the ϕ-bandwidth, ϕ_{BW}, with the normalised SFOAE level for all 20 subjects. The corresponding curves obtained from the cochlear models are shown in fig. 7.6c. Recall that for the model, the estimates of ϕ_{BW} were found to be less reliable the those of ϕ_C (fig. 6.3), leading to 'glitches' in the curve of estimated ϕ_{BW} variation, as was seen in figs. 7.6c and 7.7c for the self-suppression simulation at a probe level of 50 dB SPL. However, ignoring such glitches (which were found to be dependent on the realization of the random scattering sites) the predicted self-suppression and two-tone suppression curves together either resembled a "y" as in fig. 7.7c, or a distorted character "y" as in fig. 7.6c (the distortion being that the short limb is rotated anticlockwise). The longer limb of the "y" arises from the self-suppression experiment, and the shorter limb arises from the two-tone suppression experiment.

In order to obtain an average curve for ϕ_{BW} across both sessions, the same analysis procedure was followed as for the ϕ_C results outlined in the previous section. This led to some ϕ_{BW} estimates being judged "unrepeatable" and therefore excluded from further analysis. As would be expected, those frequency sweeps which showed repeatable trends in their ϕ_C estimates also showed repeatable trends in their ϕ_{BW} estimates. The averaged curves are shown in fig. 9.29.

The curves for each experiment were then classified as in the previous section, giving the results shown in table 9.4.

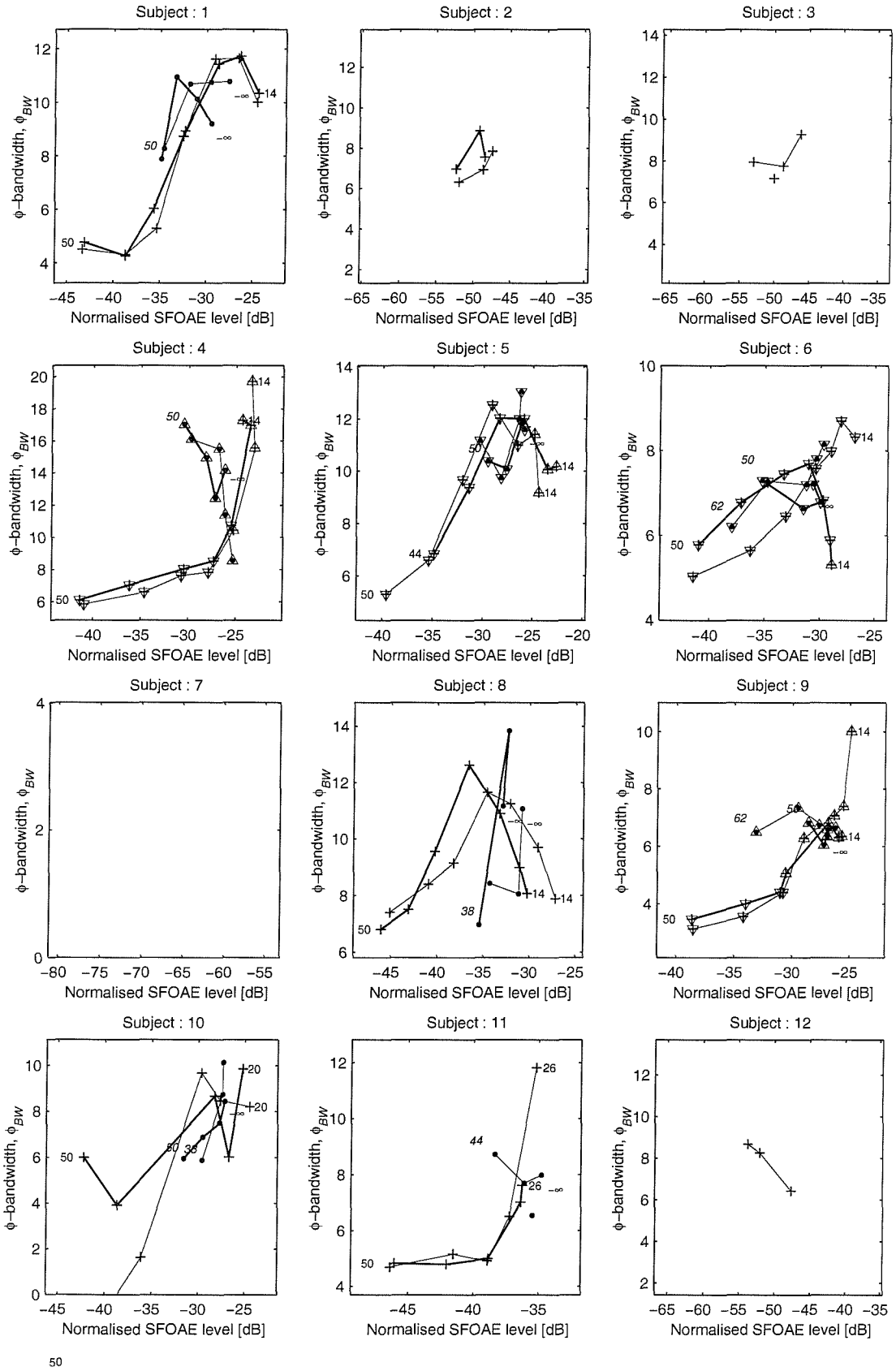


Fig 9.28 Variation of the measured SFOAE ϕ–bandwidth during both the self-suppression and the two-tone suppression experiments. Each point represents ϕ_{BW} obtained from one sweep. The horizontal axis shows the normalised SFOAE level, $p_{\Delta:\mathrm{RMS}}$. Results are shown for all 20 subjects and for both sessions. The S/EOAE rating of each sweep is also indicated. Rejected sweeps are not shown. Numbers in normal font next to selected points indicate the nominal probe level for the self-suppression experiment, for that point. Numbers in italics next to selected points indicate the suppressor level for the two-tone suppression experiment, for that point. Key shown on the following page. Continued over page.

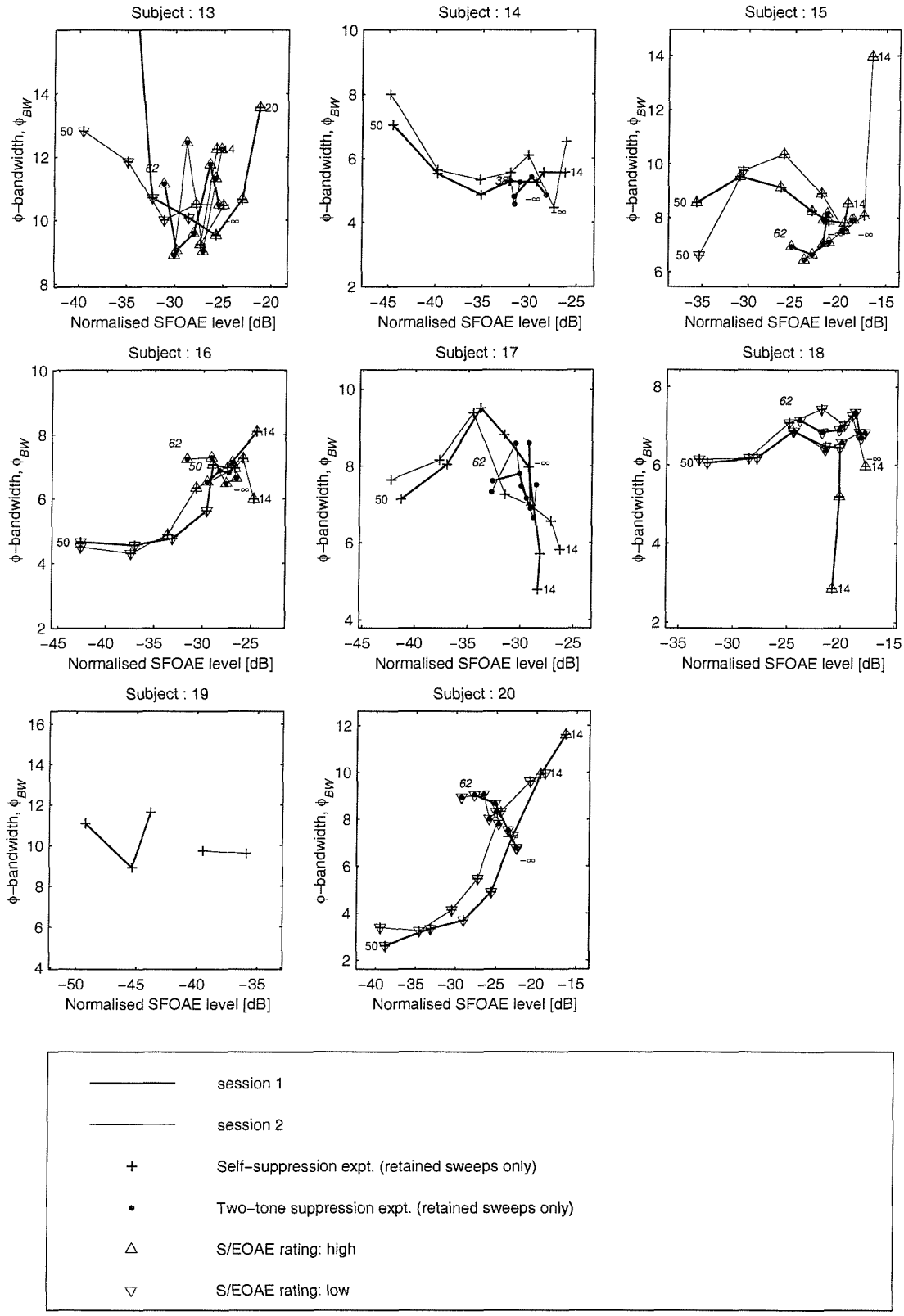


Fig 9.28 continued.

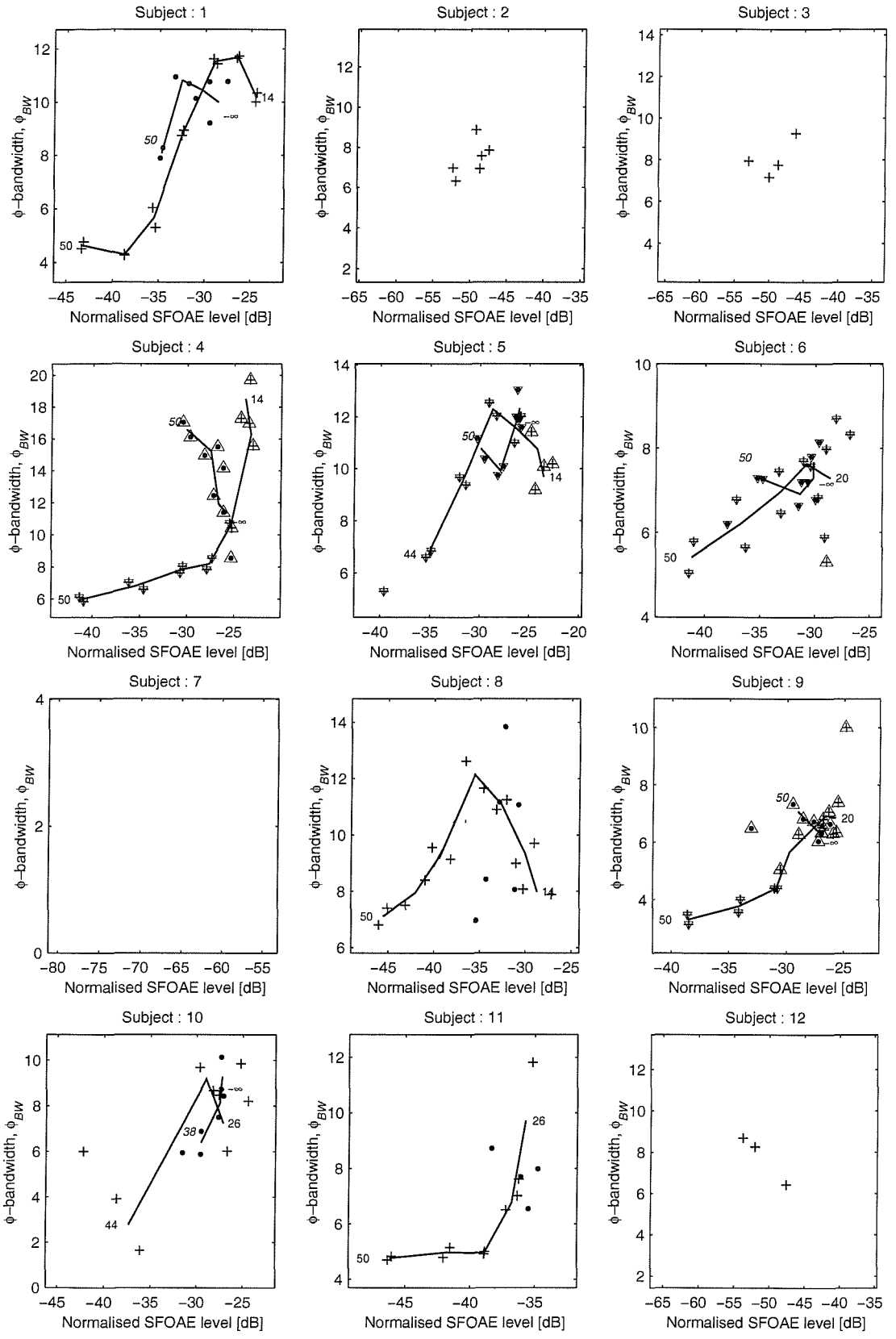


Fig 9.29 Measured ϕ–bandwidth curves defined by the average of estimates for session 1 and session 2. The variation is shown for both the self-suppression and the two-tone suppression experiments. The marked points are the actual estimates of ϕ_{BW}. The thick line passes though the average of the two estimates for session 1 and 2. Only accepted averages have been used to define the average. See text for the acceptance criteria. The horizontal axis shows the normalised SFOAE level, $p_{\Delta:\mathrm{RMS}}$. The S/EOAE rating of each sweep is also indicated. Numbers in normal font next to the start and end of a curve indicate the nominal probe levels at the end points for the self-suppression experiment. Numbers in italics next to the start and end of a curve indicate the nominal suppressor level at the end points for the two-tone suppression experiment. Key shown on the following page. Continued over page.

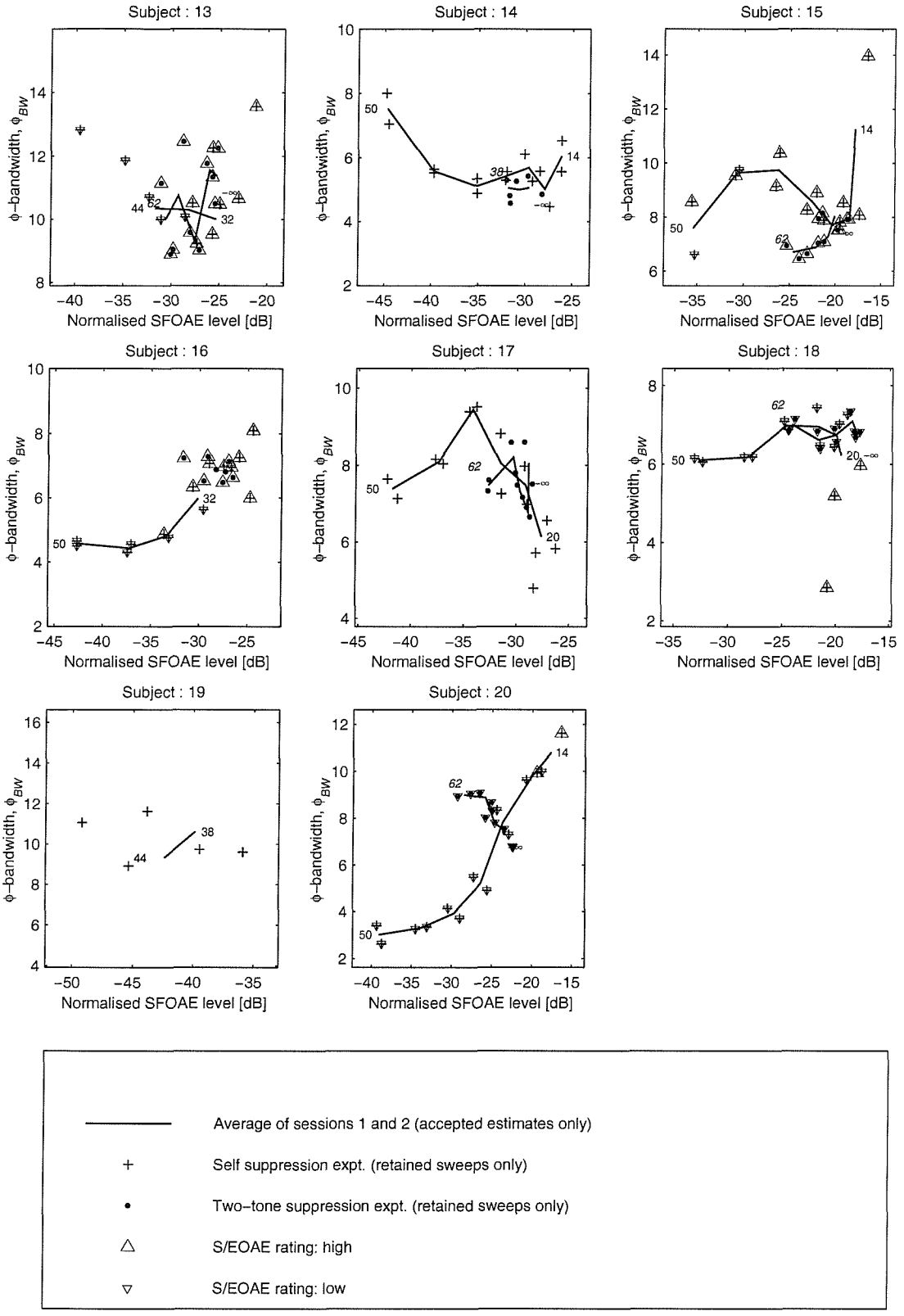


Fig 9.29 continued.

Table 9.4: Summary of results for the measured ϕ_{BW} variation with SFOAE level

Experiment	Classification of Trend	Number of Subjects	Subject Identification Numbers
Self-suppression	As model	6	4, 6, 9, 11, 16, 20
	Anomalous	9	1, 5, 8, 10, 13-15, 17, 18
	Insufficient data	5	2, 3, 7, 12, 19
Two-tone suppression	As model	2	4, 20
	Anomalous	9	1, 5, 6, 10, 13-15, 17, 18
	Insufficient data	9	2, 3, 7, 8, 9, 11, 12, 16, 19
Both experiments	As model	2	4, 20

The results show more intersubject variability and are harder to interpret than those for the ϕ_C measurements. In table 9.4, curves which show significant non-monotonicity have been classed as anomalous (despite the non-monotonicity seen in fig. 7.6c). For the self-suppression experiment, six subjects showed the trend predicted by the model. Nine self-suppression curves were classed as anomalous. The curves for subjects 1, 5, 8, 15 and 17 show a clear non-monotonicity, whereby the ϕ_{BW} estimate first increases, and then decreases as the normalised SFOAE level is reduced by the increasing probe level. A second type of anomaly is seen for subject 14 where there is a tendency for ϕ_{BW} to increase with probe level. The remaining anomalous self-suppression curves (subjects 10, 13 and 18) either show little change, or erratic changes in ϕ_{BW}.

The ϕ_{BW} curves for the two-tone suppression experiment are even harder to analyse, because of the limited numbers of accepted estimates and the limited variation achieved in the normalised SFOAE level. Consequently their was insufficient data for nine subjects. For the remaining cases, it was decided to place any "y" shaped curves (whether distorted or not) in the "as model" class. That is to say, the classification was not based on the absolute slope of the ϕ_{BW} curve for two-tone suppression, but on the slope for two-tone suppression curve relative to that for self-suppression. Only two subjects (numbers 4 and 20) clearly showed this form of ϕ_{BW} variation. The nine results which were classed as anomalous showed a variety of different trends. Subjects 1, 5 and 17 showed some non-monotonicity. Subject 15 showed a steeper ϕ_{BW} curve in the two-tone suppression case than in the self-suppression case. The remaining subjects showed a ϕ_{BW} curve in the two-tone suppression that was of a similar slope to that for the self-suppression case.

Overall, for both the self-suppression and the two-tone suppression experiments there is more intersubject variability in the results for ϕ_{BW} than was seen in the ϕ_C variation. This is expected to some extent from the results from the cochlear models given section 6.8 where the expected error in the estimates of ϕ_{BW} from the 4-parameter model was greater than those of ϕ_C. Thus, even using noise free results from the cochlear models, the estimated ϕ_{BW} variations from many realisations showed 'glitches' which resulted in non-monotonicity. As a consequence, the comparison between the measured ϕ_{BW} results and the theoretical ones is difficult, and therefore this part of the experiment does not provide a strong test of the model. Further modelling work (discussed in section 10) and experimentation may enable a stronger test based on ϕ_{BW} to be devised.

9.6.5 Variation of the Relative Bandwidth

The variation of the ϕ-bandwidth in relation to the ϕ-centre frequency is shown in fig. 9.30. If the ϕ-spectrum had an approximately constant Q factor (defined by $Q = \phi_C / \phi_{BW}$) then these curves would lie on a straight line passing through the origin, where the gradient would be inversely proportional to the Q factor. The Q factor has not been plotted directly, since the values become unreliable at small values of ϕ_{BW}. A significant intersubject variation is seen in the shape of these curves. For example, subject 15 shows a Q factor which reduces with probe level, whilst subject 20 shows the opposite trend. This suggests that use of Q-factor offers no benefits over the direct use of ϕ_{BW} as a means of characterising the width of the ϕ-spectrum. This result is in agreement with the cochlear model results shown in figs. 7.6c, 7.7c and 7.8c, which showed that a significant variation in the Q-factor arose from one realization of the scattering impedance to the next.

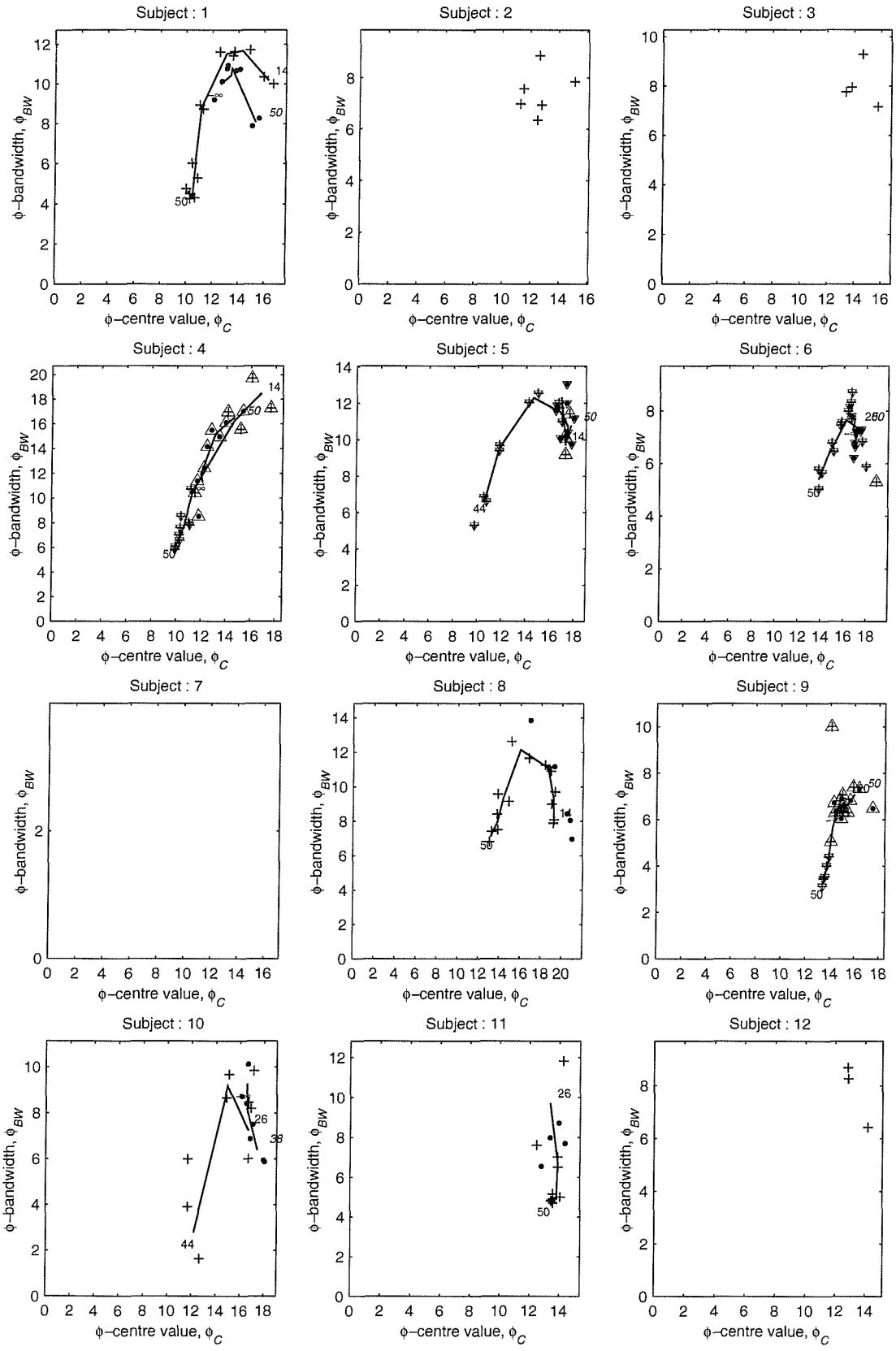


Fig 9.30 Measured ϕ-bandwidth against ϕ-centre value, using the average of estimates for session 1 and session 2. The variation is shown for both the self-suppression and the two-tone suppression experiments. The marked points are the actual estimates of ϕ_{BW} and ϕ_C . The thick line passes though the average of the two estimates for session 1 and 2. The S/EOAE rating of each sweep is also indicated. Numbers in normal font next to the start and end of a curve indicate the nominal probe levels at the end points for the self-suppression experiment. Numbers in italics next to the start and end of a curve indicate the nominal suppressor level at the end points for the two-tone suppression experiment. Key shown on the following page. Continued over page.

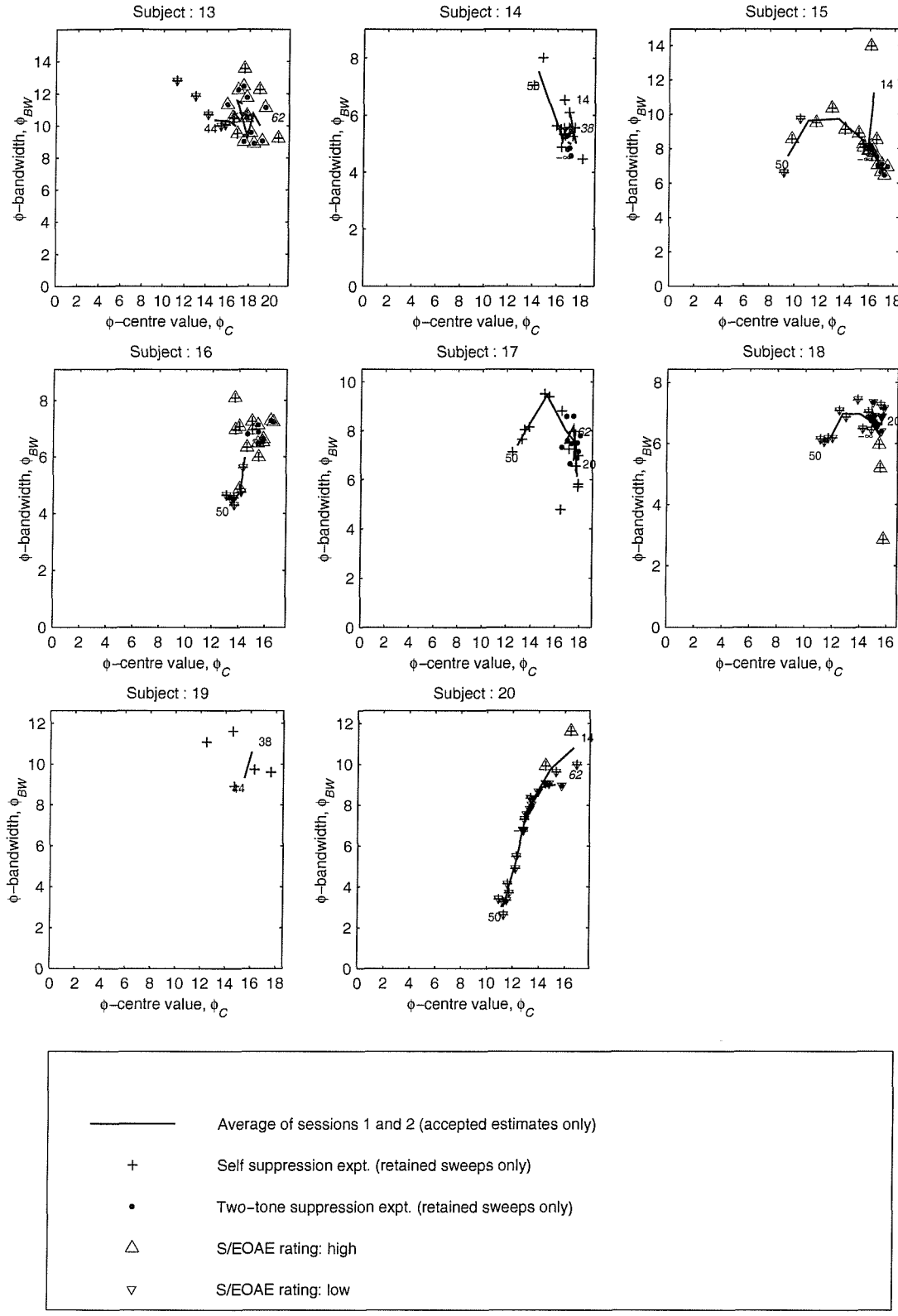


Fig 9.30 continued.

9.6.6 Variation of the Multiple Reflection Parameter, α.

Recall from section 6.6 that the value of α is related to the (frequency-averaged) product, $|\boldsymbol{rR}|$ where $\boldsymbol{r}$ and $\boldsymbol{R}$ are the cochlear basal and apical reflection coefficients respectively. Also, the value β is related to the frequency averaged product $|\boldsymbol{gR}|$ where $\boldsymbol{g}$ characterises transmission through the middle ear, outer ear and probe. The value of β is also virtually identical to the value of the RMS normalised SFOAE pressure. The effect of the probe or suppressor level is to reduce the magnitude of $\boldsymbol{R}$, leaving $\boldsymbol{r}$ and $\boldsymbol{g}$ unaltered. Therefore it is expected that α will be proportional to β. This has been tested by plotting α against β for the measured data (fig. 9.31). For comparison, results from the cochlear model predictions are shown in figs. 7.6d and 7.7d. In general, the results bear out the predictions, as is most clearly seen for subjects 4, 5, 6, 9, 14, 15, 16, 17, 18 and 20. Recall also that the studies of the performance of the 4-parameter model revealed that errors in the estimates of α were large, even for noise free data. Therefore a significant departure of the measured results from ideal proportionality (as seen in fig. 9.31) was expected.

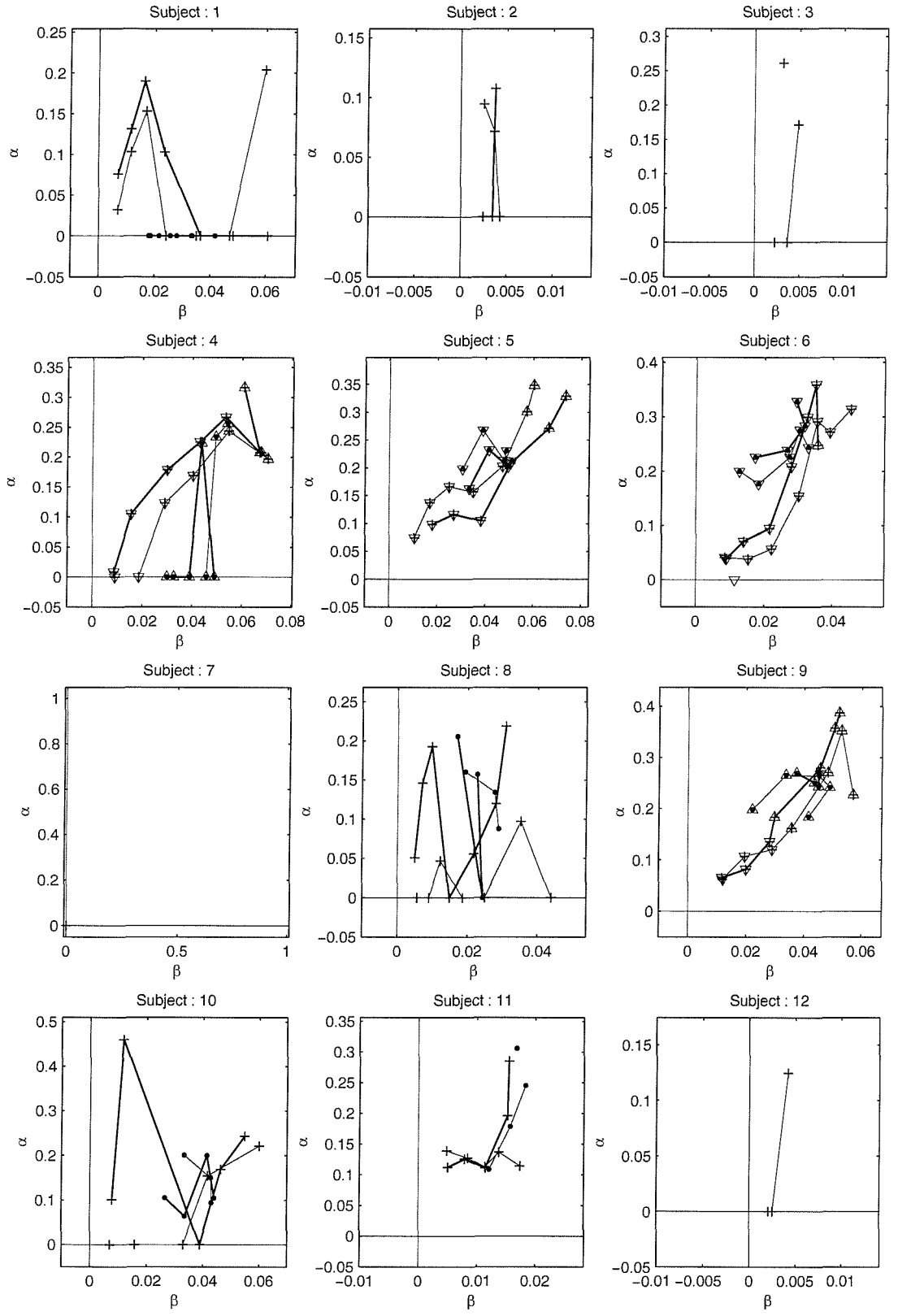


Fig 9.31 Variation of the measured α parameter during both the self-suppression and the two-tone suppression experiments. Each point represents α, which quantifies the strength of multiple reflections, obtained from one frequency sweep. The horizontal axis shows the β parameter, which is approximately equal to $p_{\Delta:\,\mathrm{RMS}}$. Results are shown for all 20 subjects and for both sessions. The S/EOAE rating of each sweep is also indicated. Rejected sweeps are not shown. Key shown on the following page.

Continued over page.

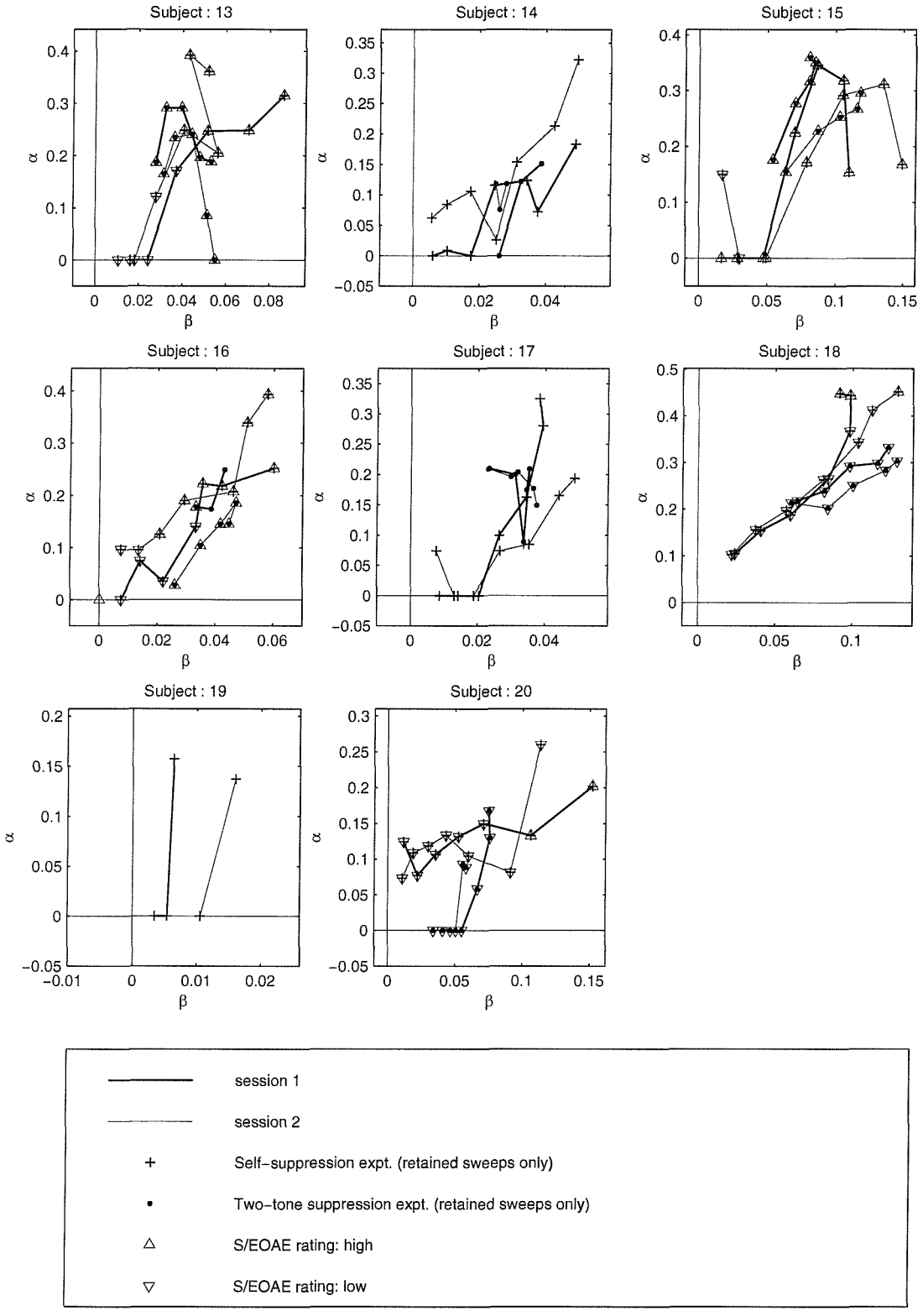


Fig 9.31 continued.

9.7 Overall Comparison of Modelling and Experimental Results

The results allow four separate comparisons between measurements and theory to be made. These are the estimates of the ϕ_C and ϕ_{BW} variations for both the self-suppression and two-tone suppression experiments.

The results for ϕ_C for the self-suppression gave the most consistent pattern of results across subjects. For nearly all subjects where the SFOAEs were strong enough to be measured accurately for several level series, the measured variation of ϕ_C showed a striking similarity to the theoretical results obtained from cochlear models. The results of the two-tone suppression experiment were less clear, partly because of the limitations of this experiment, discussed later. However, in general the results for ϕ_C for the two-tone suppression experiment also agreed with theoretical predictions. Thus, in general, the experimental results showed the characteristic y-shaped variation in ϕ_C predicted by the models. Though two of the subjects showed results that deviated slightly from the model predictions, these anomalies were weak, and could be attributed to the random variation which is inherent in the SFOAE generation mechanism proposed by Shera and Zweig. Thus, the measured ϕ_C variations have been interpreted as offering strong experimental support for the cochlear models. Furthermore, this conclusion appeared to be valid across all subjects where SFOAEs were strong enough to allow comparisons with model predictions to be made.

The precise details of the y-shaped results varied significantly across subjects. It is speculated that this variation is due to inter-subject differences in the cochlear amplifier characteristics, leading to differences in the way in which the TW shape is altered during acoustic suppression.

The results for ϕ_{BW} are harder to interpret for two reasons. Firstly the results of the cochlear modelling study showed that the errors in the 4-parameter model estimates of ϕ_{BW} were significant, leading to glitches in the ϕ_{BW} curves. Secondly, the ϕ_{BW} experimental results showed a larger intersubject variability than did the ϕ_C results (as would be expected from the modelling results). Overall, neither the self-suppression ϕ_{BW} results nor the two-tone suppression ϕ_{BW} results were clear enough to be safely interpreted as either supporting or contradicting the cochlear model. Therefore this part of the experiment does not provide a good test of theory.

Overall the experimental results show that changes in OAE periodicity, as characterised by the ϕ_C variable, can be detected during both self-suppression and two-tone

suppression. Furthermore, these changes are in agreement with cochlear models based on Shera and Zweig's theory of OAE generation combined with the model of self-suppression and two-tone suppression developed by Kanis and de Boer. This suggest that the cochlear models are accurately representing the main features of the OAE generation mechanism and of OAE suppression, and lends some support to both Shera and Zweig's theory and to Kanis and de Boer's model of nonlinearity. The results are not in agreement with a cochlear models containing Strube's spatially periodic scattering impedance.

9.7.1 A Possible Alternative Interpretation

Talmadge *et al.* (2000) have shown that some nonlinear cochlear models predict SFOAE spectra with two components: a nonlinear component and a 'roughness' or scattering component. In this thesis, only the scattering component has been included in the model predictions. As a consequence of the definition of the reflectionless cochlea (section 5.4), even the nonlinear models in this thesis show no nonlinear SFOAE component. Thus the y-shaped variation seen in the model results is due entirely to the interaction of the TW shape with the scattering impedance. However, in the following discussion, the possibility of a nonlinear effect on the experimental data is addressed.

In the model of Talmage *et al.* (2000) the nonlinear component varies only very slowly with frequency (i.e., it has a very high periodicity) and is generally of smaller amplitude to the scattering component which has been the focus of this thesis. Thus the total SFOAE spectrum is usually dominated by the scattering component. However, as the stimulus level increases the nonlinear component increases whilst the scattering component reduces. When a stimulus level is reached at which the nonlinear component becomes dominant, then the fine structure in the total SFOAE spectrum virtually disappears being replaced by the much coarser (i.e., higher periodicity) spectral structure of the nonlinear component (Talmadge *et al.*, 2000, fig 4). This would correspond to a ϕ-centre value, ϕ_C, approaching zero. It could be speculated that, in the transition region between these two regimes, the variation of ϕ_C with stimulus level might appear qualitatively similar to variation seen experimentally.

The following arguments are presented to support the interpretation given in section 9.7, that the y-shaped variation is dominated by the scattering SFOAE component, rather than by transition from scattering to nonlinear components. Firstly, no sign of the nonlinear SFOAE component was seen when comparing the suppressed and unsuppressed ear canal pressure spectra (section 9.1). Secondly, the final measured value of ϕ_C in

fig 9.27 is typically 10. This is much higher than the value of ϕ_C due to the nonlinear component alone (which is less than 1 in the frequency range of interest in Talmadge *et al.*, 2000, figs 2, 4 and 6). This shows that the final SFOAE is still dominated by the scattering component. Thirdly, the variation in ϕ_C with stimulus level follows a smooth, continuous line. If a transition phenomenon were strongly influencing the results, it might be expected that a more sudden drop in the value of ϕ_C with stimulus level would be seen. Finally, the striking resemblence of the experimentally measured ϕ_C variations to those obtained theoretically argues against the involvement of nonlinear SFOAE components. Nevertheless, without further modelling studies, which are outside the scope of this thesis, it is difficult to completely rule out any influence from nonlinear SFOAE components. Such further modelling studies are left as recommended future areas of investigation.

9.8 Implications of the Results

Within the framework of Kemp's reflection hypothesis, the main questions that have a bearing on the results obtained in this thesis are:

- What is the spatial form of the scattering impedance in the cochlea?
- What is the nonlinear variation in the TW response during self-suppression and two-tone suppression?
- What is the relationship between SFOAEs, the TW response and the scattering impedance?
- What is an appropriate signal processing method for characterising SFOAE periodicity?

In this thesis, the primary aim was to acquire experimental evidence which would shed light on the form of the scattering impedance. However, the answers to the remaining questions listed above are still uncertain. Therefore the assumptions that have been made in answering these questions may influence any conclusions drawn about the scattering impedance.

In this respect the interpretation of the self-suppression experiment given in the previous section is less open to question than that of the two-tone suppression experiment. This is because the broadening of the TW envelope with increasing probe level is well established from direct measurements of the BM mechanical response (Cooper and Rhode, 1992). Furthermore, the modelling of self-suppression is more straightforward than two-tone suppression. In fact, the results of the cochlear models presented here show that the

changes in TW shape during self-suppression predicted by the nonlinear model are similar to those predicted by linear models in which the overall cochlear amplifier gain is reduced. This has been found for two quite different (linear) formulations of the cochlear amplifier. For the random scattering impedance models, the resulting changes in ϕ_C and ϕ_{BW} were seen for both of these linear models as well as for the nonlinear model. This suggests that the general result is not dependent on the details of the cochlear model.

In contrast, several factors make the interpretation of the two-tone suppression experiment more questionable. The changes in TW shape in two-tone suppression are generally less well established than those in self-suppression. Furthermore, model predictions of two-tone suppression necessarily require a nonlinear formulation: there is no simple way of checking the nonlinear model results against a linear model. Such nonlinear formulations introduce another set of assumptions, over and above those required for the linear models. Due to time constraints, only the nonlinear model due to Kanis and de Boer was implemented. Therefore there is a danger that the results depend on the details of this particular formulation. Finally, the conclusions that have been drawn are based on a single suppressor-to-probe frequency ratio, and at a single probe level.

The self-suppression results are in agreement with the result reported by Kemp and Brown (1983), that increasing the probe level gave a reduction in group delay. They are also in agreement with the reported results of a reduction of TEOAE latency with probe level (Neely *et al.*, 1988). No reported results have been found which either directly confirm or contradict the two-tone suppression results.

The experimental results have been interpreted as providing support for Shera and Zweig's theory in which the CP scattering impedance shows a fine-grained, random spatial variation. This requires further qualification. Two main classes of scattering model have been used in the simulations: random scattering and periodic scattering models. The measured ϕ_C variations clearly confirm the predictions of the random scattering models but contradict those of the periodic scattering models. However, it may be argued that the periodic scattering models used were unrealistic, since the resulting SFOAE spectra show near perfect periodicity (section 7.2). This can be seen by considering their predicted ϕ_{BW} values, which are much lower than measured values (figs. 7.3 and 9.28). Thus such models could have been discounted purely on the measured value of ϕ_{BW} for a single frequency sweep. Although this problem was partially addressed in section 5.5.1, where more realistic SFOAE results were presented from models which have a narrow-band, rather than a periodic, scattering impedance, these models were not included in the self-suppression or two-tone suppression simulations. However, it is possible to predict the

results that would have been obtained by considering Zweig and Shera's phenomenological theory, as illustrated in figs. 4.4 and 4.5. For a scattering impedance, the scattering potential in panel (d) would have a narrow band bandpass shape in the ϕ-domain, somewhere between the random broadband function seen in fig. 4.4d and the sharp peak seen in fig. 4.5d. Consider multiplying this bandpass shape with the TW pulse in the ϕ-domain. It can be seen that the resulting SFOAE periodicity may or may not depend strongly on the TW shape, depending on the relative positions in the ϕ-domain of the scattering potential and the TW function. We might then class those cochlear models which predict SFOAE spectra whose periodicity is dominated by the peak of the TW function as representing Shera and Zweig's model, and those which predict SFOAE spectra whose periodicity is dominated by the peak in the scattering potential as representing Strube's model. The measured results indicate that significant changes in ϕ_C occur during suppression, which strongly suggests that the SFOAE periodicity is not dominated by the spatial periodicity of the scattering impedance. In this sense, the results contradict Strube's model.

However, this result does not necessarily imply that a perfectly random scattering impedance was actually present in the tested ears. Other spatial variations which have not been simulated might produce similar results. For example, perhaps the fine-grained inhomogeneities could be replaced by fewer point inhomogeneities rather like the models discussed in section. 3.3.1, except with greater numbers of reflection sites, and with a random distribution. This could be described as a sparsely distributed, coarse grained random distribution. Or a BM with some scattering regions interspersed with smooth regions might produce similar results, depending on the size and location of these regions. These two suggestions are made simply to illustrate that there are many possible variants of the random scattering impedance which have not been simulated. Further modelling and experimentation would allow the models to be refined.

Overall the experimental results for the self-suppression experiment provide strong support for the model of OAE generation based on Shera and Zweig's theory. The results of the two-tone suppression experiment appear to support both Shera and Zweig's theory and the two-tone suppression model developed by Kanis and de Boer. However, further modelling and experimentation in this area is desirable.

The results do not refute the hypothesis that Shera and Zweig's theory is generally applicable to all normally hearing subjects. The few anomalies that were seen are readily accommodated by the random nature of the proposed OAE generation mechanism.

If the interpretation outlined above is correct, then changes in the shape of the TW function manifest themselves as measurable changes in the characteristics of the SFOAEs. Therefore it may be possible to extract potentially interesting information about TW. This may provide a method of linking the results obtained for SFOAEs with measures such as TEOAEs, or with psychophysical measures of the auditory filter. It may also be possible to study further the nonlinear behaviour of the TW using SFOAEs.

Unfortunately, however, the information about the TW is partially obscured by the random nature of the SFOAE, which according to theory arises from the random arrangement of inhomogeneities in the cochlea. This means, for example, that models predict different values of the SFOAE periodicity (or of ϕ_C) from different realizations of the scattering impedance, even when the primary TW response (and therefore the auditory filter width) is held constant (section 7.3). This suggests that accurately measuring auditory filter bandwidth using SFOAEs or TEOAEs may be problematic. This conclusion is, however, based on the use of the 4-parameter model to measure periodicity. Better signal processing methods may reveal this information more clearly.

PART IV: CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE WORK

10. Conclusions

The aim of this investigation was to test theories of the mechanism of OAE generation against experimental evidence. This aim was achieved by examining the variation of the periodicity of the SFOAE during self-suppression and two-tone suppression. This investigation falls into three main parts.

In the first part, a macromechanical cochlear model has been developed which allows predictions of SFOAEs to be made. This is a simple longwave model, incorporating models both of the cochlear amplifier and of reflection sites on the basilar membrane for the cochlear TW. These reflection sites were modelled by a scattering impedance implemented as spatial variations in the basilar membrane point mechanical impedance. In one group of cochlear models, the reflection sites were modelled as a random component in the spatial variation of the basilar membrane impedance, as proposed in Shera and Zweig's theory. In variants of these models, this random spatial variation is replaced by a periodic variation, as proposed by Strube. Using these two groups of models, SFOAEs predicted by Shera and Zweig's theory can be compared to those predicted by Strube's model. The models also include cochlear amplifier nonlinearity, based on the quasilinear implementation of Kanis and de Boer, which allows the effects of self-suppression and two tone suppression on SFOAEs to be simulated. In an additional model variant, self-suppression was crudely simulated by an overall reduction in the cochlear amplifier gain of a linear model.

Various model parameter studies have been performed to examine the sensitivity of the predicted SFOAEs to the modelling assumptions. The sensitivity to changes in the middle-ear model, to the scattering impedance, and to the cochlear amplifier characteristics were studied. It was found that the major results reported in this thesis are all insensitive to these changes, suggesting that they are generally applicable to a large class of cochlear models.

In the second part of this thesis, a method has successfully been developed for accurately determining the periodicity of an SFOAE spectrum. Because of the random nature of the scattering impedance in Shera and Zweig's theory, the SFOAE spectrum has been treated as a single realization of an underlying random process. The random nature of the SFOAE makes the definition of spectral periodicity problematic. A method has been developed for quantifying the periodicity of the ripple pattern seen in a single SFOAE

frequency spectrum (whether measured or predicted). This method, known as the 4-parameter model, has been designed to extract parameters that characterise the random process arising from Shera and Zweig's theory of SFOAE generation. Physically these parameters relate to the shape of the cochlear TW, and to the extent and degree of multiple reflection in the cochlea. One of the parameters, called the ϕ-centre frequency and denoted ϕ_C, relates directly to the periodicity of the SFOAE ripple pattern. This effectively provides a definition of periodicity. A second parameter, called the ϕ-bandwidth and denoted ϕ_{BW}, quantifies the spread of periodicities in the SFOAE spectrum. In cochlear modelling studies, it was found that the 4-parameter model enables small changes in the periodicity of the SFOAE arising in self-suppression and two-tone suppression studies to be detected. The 4-parameter model has been applied to the predicted SFOAEs from the cochlear models during both self-suppression and two-tone suppression.

In the third part of this thesis, experiments were carried out on 20 normally hearing subjects to measure the variation of the SFOAE periodicity during self-suppression and two-tone suppression. The two-tone suppression was carried out with a high-side suppressor tone whose frequency was 1.3 times that of the probe tone. A subset of the subjects was created containing only those subjects who showed accurately measurable SFOAEs over a range of suppression levels. The 4-parameter model was then used to compare predicted and measured variations in ϕ_C and ϕ_{BW} which characterise the periodicity.

The following conclusions have been drawn from the investigation.

1. Those cochlear models that are based on Shera and Zweig's theory of OAE generation produce realistic looking SFOAE predictions. Two features of the model that are necessary for predicting realistic SFOAEs are first, a random spatial variation in the basilar membrane impedance and second, a cochlear amplifier giving active TW amplification. These results are quite insensitive to the major assumptions in the model, in that different formulations of the cochlear amplifier and of the random scattering impedance do not lead to qualitative differences in SFOAEs. These results confirm Zweig and Shera's results (1995).
2. The SFOAEs predicted by cochlear models based on Shera and Zweig's theory show distinctive changes in the ϕ-centre value, ϕ_C, and ϕ-bandwidth, ϕ_{BW}, during both self-suppression and two-tone suppression. Specifically, during self-suppression the value of ϕ_C reduces as the normalised SFOAE level reduced. However, during high-side,

two-tone suppression the value of ϕ_C increases as the normalised SFOAE level reduces. This variation resembles a y-shape when ϕ_C is plotted against the normalised SFOAE level. In general ϕ_{BW} shows qualitatively similar trends to those of ϕ_C during both self-suppression and two-tone suppression. However, the estimates of ϕ_{BW} are less reliable than those of ϕ_C (i.e., they showed a greater variation across an ensemble of models, each with a different realization of the random scattering impedance). In some cases, for a particular realization of the random scattering impedance, the trend in ϕ_{BW} with the normalised SFOAE level is not monotonic. Thus overall the models shows that ϕ_{BW} is less useful than ϕ_C as a measure of either periodicity or TW shape.

3. The 4-parameter model provides a useful stochastic model of an SFOAE frequency spectrum obtained at constant stimulus level. This gives a measure of SFOAE periodicity which can be related to physical quantities in Shera and Zweig's theory of SFOAE generation.
4. Experiments show that measured SFOAE ϕ_C and ϕ_{BW} are altered both during self-suppression and two-tone suppression.
5. In the subset of subjects where SFOAEs could be measured with sufficient accuracy, the variations of the measured ϕ_C show a similar y-shaped variation to that seen in the predictions of the cochlear models based on Shera and Zweig's theory of OAE generation. These results are interpreted as providing strong supporting evidence for this theory.
6. Experimental results for ϕ_{BW} were harder to interpret as the observed trends were often not monotonic. It is concluded that these results are not clear enough to be safely interpreted as either supporting or contradicting the cochlear model. Further modelling and experimental work is recommended in this area.
7. Overall, it is concluded that the measured variation in SFOAE periodicity provides strong support for Shera and Zweig's theory of OAE generation, thereby suggesting a fundamental link between SFOAE periodicity and the shape of the cochlear TW.

11. Recommendations for Future Work

Recommendations are made here for work in three areas: experimental, cochlear modelling, and signal processing.

11.1 Improvements to the Experimental Procedure

One of the main limitations of the experiment was the difficulty in obtaining sufficient SFOAE data to allow trends to be assessed. Despite choosing subjects with normal hearing (generally better than 10 dB HL) and recording for about three hours in total, the results from about a quarter of the subjects were too noisy to be useable. One practical limitation of the recording method was that SFOAEs were not calculated at measurement time. If this were rectified, the recording system could adaptively vary the number of averages to achieve a desired SNR. The experimental procedure could be modified to measure at fewer levels, but with greater accuracy in those subjects with weaker SFOAEs.

11.2 Further Experiments

The following further experiments are suggested:

1. a more comprehensive study of the effects reported here in which self-suppression and two-tone suppression of SFOAEs are studied in more depth,
2. a characterisation of level effects and suppression effects in TEOAEs,
3. a study of the effects of ototoxic drugs or contralateral suppression on the SFOAE ϕ-spectrum,
4. a study of the relationships between the ϕ-spectrum and the psychophysical measures of the auditory filter bandwidth,
5. experiments specifically designed to separate possible nonlinear SFOAE components from the scattering SFOAE component.

These proposals are outlined below.

Firstly, a more comprehensive study of SFOAEs in self-suppression and two-tone suppression could be undertaken. Though the results of the self-suppression experiment reported here provided an adequate test of the model, the two-tone suppression experiment was less successful, owing partly to the limited level of suppression that was obtained (typically 7 dB). A lower suppressor-to-probe frequency ratio (say 1.1 or 1.2) would give a

higher degree of suppression. A study involving a range of suppressor to probe frequency ratios would provide further tests of the nonlinear elements of cochlear models. Studies could also explore the other experimental parameters which were fixed in the experiments reported here: the probe frequency range and the probe level in two-tone suppression experiments. Additionally, a further investigation of the anomalous results found for ϕ_{BW} would also be interesting, though this should follow a more thorough modelling study.

A second area which might be fruitfully explored is the characterisation of TEOAEs. Cochlear theory predicts, and measurements confirm, that the SFOAE frequency spectrum is similar to the Fourier transform of a TEOAE waveform. In fact the following argument shows that a TEOAE can be thought of a suppressed SFOAE. For simplicity, consider the input stimulus to a TEOAE as a periodic train of clicks, with a click interval greater than the TEOAE duration. This can be decomposed into the sum of frequency components of roughly equal strength comprising the Fourier series expansion. Each frequency component in the measured TEOAE can therefore be thought of as an SFOAE, suppressed by multiple suppressor tones. More generally, when the click train is not periodic, it can be seen that the TEOAE frequency spectrum is a particular class of suppressed SFOAE frequency spectra: a SFOAE frequency spectrum measured in the presence of a broad band suppressor. An obvious experiment to perform would be to measure TEOAEs at various click levels, and to analyse their frequency spectra using the 4-parameter model. The effect of various additional suppressor stimuli on the TEOAE could also be studied. Since in TEOAE level series the effects of self-suppression and multi-tone suppression are combined, it is not possible to predict how the TEOAE spectra should vary, without performing further modelling studies.

In the third area proposed above, experiments could investigate other methods of altering the TW, such as ototoxic drugs or contralateral acoustic suppression. Measuring the SFOAE frequency spectrum, and using the 4-parameter model to extract the periodicity (or ϕ-spectrum) may provide information about changes in the TW shape.

One problem with using ototoxic drugs, such as aspirin, is that they take days to have any effect. In the experiments reported here, it was found that the ϕ-spectrum can be altered simply by removing and refitting the OAE probe (perhaps because this alters the parameters denoted g and r in section 4.10). Day-to-day changes in the middle ear would also be a problem. The acoustic effects investigated here do not suffer from these shortcomings, since the OAE probe can be left in place for the duration of the experiment. Day-to-day changes in the ϕ-spectrum are seen in the session-to-session variations of the results reported here.

A problem with contralateral suppression is that the effect is small (typically giving changes of only a few dB). Nevertheless there may be some benefit in investigating this effect.

In the fourth area proposed, experiments could look for a relationship between the SFOAE periodicity and psychophysical measures of the auditory filter. If, as is generally believed, the width of the auditory filter is determined primarily by the width of the mechanical tuning curve, then both the psychophysical measure, and the ϕ-spectrum will be determined by the shape of the TW. As a consequence, it might be thought that subjects with narrower tuning curves measured psychophysically would also show smaller periodicities (or higher ϕ_C and ϕ_{BW} values). However, such a simple relationship does not necessarily follow from the cochlear models. This is because the two measures are different, single number characterisations of the TW shape, which is a complex curve. Inter-subject differences in TW shape would confound the results. Furthermore, the TW shape is dependent on the acoustic stimulus. The psychophysical measures are usually performed at high overall levels (taking both probe and masker into account), and with quite a complex acoustic stimulus (such as a pure tone plus a notched noise masker). Now, ϕ_C has been found to be highly dependent on both probe and suppressor (or masker) level. Also, this dependency varies across subjects, such that a high ϕ_C at one level is not necessarily a good predictor of the ϕ_C at a different level. Therefore any experiment would have to account for these effects in order to relate the two measures. It might be possible to use nonlinear cochlear models to help assess the importance of these factors.

In the final area, an experimental investigation of the influence of nonlinear SFOAEs reported by Talmadge *et al.* (2000) may be worthwhile. This should be undertaken following further modelling work in the area which might help establish the characteristics of these components.

11.3 Development of Cochlear Models and Signal Processing Techniques

This thesis has relied on both macromechanical models of the cochlea and on the 4-parameter model for extracting information from measured SFOAEs. Both of these areas could be further developed.

Many properties of the cochlea are still poorly understood. Consequently there is much scope for further model development. In the absence of reliable measurements in the real cochlea, studies could be undertaken to assess the sensitivity of cochlear model

predictions to changes in those parameters which are still very uncertain. Aspects which could be addressed are the effects on the predicted ϕ-spectrum of:

- nonlinear cochlear amplifier effects,
- the form of the basilar membrane inhomogeneity,
- scaling symmetry in the cochlea.

The most important of these aspects is probably the modelling of the cochlear amplifier. One simple extension to the modelling work reported here would be to modify the quasilinear model to include the Zweig cochlear amplifier (1991), rather than the Neely and Kim (1986) cochlear amplifier. The simulation of the self-suppression and two-tone suppression cases could then be repeated and the results compared with experimental data. This would reduce the likelihood of making misleading predictions, which could be sensitive to the choice of cochlear amplifier model. More generally, Kanis and de Boer's formulation of cochlear nonlinearity contains many assumptions and simplifications. It may be interesting to investigate alternative formulations of this nonlinearity.

A second aspect of the models that may be investigated is that of TW reflection sites. The physiological basis for TW reflection is still uncertain. Preliminary investigations using the cochlear models (section 4.6) have shown that the form of the basilar membrane inhomogeneity (e.g., whether the inhomogeneity is in the basilar membrane stiffness or damping) has a significant effect on the model predictions of SFOAEs. The effects of these formulations on the ϕ-spectrum could be studied more comprehensively. In addition, the various forms of spatial variation of scattering impedance discussed in section 9.8 could be further investigated. For example, a narrow band spatial variation, a coarser grained random distribution, or an intermittent random spatial distribution could all be modelled quite easily.

Finally in the area of cochlear modelling, improvements to the modelling of scaling symmetry in the cochlear model might allow a better prediction of the variation of periodicity with frequency. The justification for converting the SFOAE frequency spectrum to a log-frequency axis (or η-domain) is to straighten out the phase spectrum. Viewed from the time domain, this is equivalent to reversing the effect of frequency dispersion in the cochlea. However, the logarithmic transformation is only strictly accurate when perfect scaling symmetry holds. A more accurate transformation, based on some average cochlear properties, should lead to a better characterisation of the SFOAE response.

The stochastic model of the SFOAE frequency signal is another area where further improvements can perhaps be made. This is an important area since it determines the

accuracy with which potentially useful information about the cochlea can be obtained from SFOAE measurements. The basic 4-parameter model adopted in this thesis is a very simply one: a stationary, gaussian white noise signal is first passed through a Butterworth filter, and then a transformation is applied to account for multiple reflections. Further work could look at:

- non-gaussian signals,
- non-stationary signals,
- different filter transfer functions,
- different treatments of multiple reflections.

Additional parameters might need to be added to account for some of these effects. In addition, improvements to the estimation algorithm may be possible. This work could either be done in conjunction with cochlear models, or it could be based entirely on measured signals.

12. References

Ables, J. G. (1978) 'Maximum entropy spectral analysis' in D. G. Childers (Ed.), *Modern Spectrum Analysis*, IEEE Press Selected Reprint Series, John Wiley and Sons, New York, pp. 23-33.

Allen, J. B. (1988) 'Is basilar membrane tuning the same as neural tuning -- Where do we stand?' in J. P. Wilson and D. T. Kemp (Ed.), *Cochlear mechanisms - structure, function, and models*, (NATO ASI Series), Plenum Press, New York, pp. 453-463.

Allen, J. B. (1990) 'Modeling the noise damaged cochlea' in P. Dallos, C. D. Geisler, J. W. Matthew, M. A. Ruggero, C. R. Steele (Ed.), *The Mechanics and Biophysics of Hearing*, Springer-Verlag, Berlin, pp. 324-332.

Avan, P., Wit, H. P., Guitton, M., Mom, T. and Bonfils, P. (2000) 'On the spectral periodicity of transient-evoked otoacoustic emissions from normal and damaged cochleas' J. Acoust. Soc. Am. 108, pp. 1117-1127.

Bendat, J. S. and Piersol, A. G. (1966) *Measurement and analysis of random data*, John Wiley and Sons, New York.

Brass, D. and Kemp, D. T. (1993) 'Suppression of stimulus frequency otoacoustic emissions' J. Acoust. Soc. Am. 93, pp. 920-939.

Brown, A. M., Williams, P. M. and Gaskill, S. A. (1993) 'The effect of aspirin on cochlear mechanical tuning' J. Acoust. Soc. Am. 93, pp. 3298-3307.

Burg, J. P. (1978a) 'Maximum Entropy Spectral Analysis' in D. G. Childers (Ed.), *Modern Spectrum Analysis*, IEEE Press Selected Reprint Series, John Wiley and Sons, New York, pp. 34-41.

Burg, J. P. (1978b) 'A new analysis technique for time series data' in D. G. Childers (Ed.), *Modern Spectrum Analysis*, IEEE Press Selected Reprint Series, John Wiley and Sons, New York, pp. 42-48.

Collet, L., Veuillet, E., Moulin, A., Morlet, T., Giraud, A. L., Michyl, C., and Chéry-Croze, S. (1994) 'Contralateral auditory stimulation and otoacoustic emissions: a review of basic data in humans' Br. J. Audiol. 28, pp. 213-218.

Cooper, N. P. and Rhode, W. S. (1992) 'Basilar membrane mechanics in the hook region of cat and guinea pig cochleae: Sharp tuning and nonlinearity in the absence of baseline position shifts' Hear. Res. 63, pp. 163-190.

Dallmayr, C. (1987) 'Stationary and dynamical properties of simultaneous evoked otoacoustic emissions (SEOAE)' Acustica 63, pp. 223-255.

Dallos, P. (1996) 'Overview: Cochlear Neurobiology' in P. Dallos, A. N. Popper, R. R. Fay (Ed.), *The Cochlea*, Springer-Verlag, New York, pp. 1-43.

Dallos, P., Popper, A. N., Fay, R. R. (Ed.) (1996) *The Cochlea*, Springer-Verlag, New York.

de Boer, E. (1980) 'Dimensionality of cochlear waves' in G. van den Brink, F. A. Bilsen (Ed.), *Psychophysical, Physiological and Behavioural Studies in Hearing*, Proceedings of the 5th International Symposium on Hearing, Delft University Press, The Netherlands, pp. 34-42., pp. 2-6.

de Boer, E. (1983) 'No sharpening - a challenge for cochlear mechanics' J. Acoust. Soc. Am. 83, pp. 567-573.

de Boer, E. (1993) 'Some like it active' in H. Duifhuis, J. W. Horst, P. van Dijk, S. M. van Netten (Ed.), *Biophysics of Hair-Cell Sensory Systems*, World Scientific, Singapore, pp. 3-22.

de Boer, E. (1995) 'On equivalence of active models of the cochlea' J. Acoust. Soc. Am. 98, pp. 1400-1409.

de Boer, E. (1996) 'Mechanics of the Cochlea: Modeling Efforts' in P. Dallos, A. N. Popper, R. R. Fay (Ed.), *The Cochlea*, Springer-Verlag, New York, pp. 258-317.

de Boer, E., Kaernbach, C., König, P., Schillen, T. (1986) 'Forward and reverse waves in the one-dimensional model of the cochlea', Hear. Res. 23, pp. 1-7.

de Boer, E. and MacKay, R. (1980) 'Reflections on reflections', J. Acoust. Soc. Am. 67, pp. 883-889.

de Boer, E., Kaernbach, C., König, P., Schillen, T. (1986) 'Forward and reverse waves in the one-dimensional model of the cochlea', Hear. Res. 23, pp. 1-7.

Deutsch, R. (1965) *Estimation Theory*, Prentice-Hall, Englewood Cliffs, New Jersey.

Eggermont, J. J. (1979) 'Compound action potentials: Tuning curves and delay times' in M. Hoke and E. de Boer (Ed.), *Models of the auditory system and related signal processing techniques*, Scand. Audiol. Suppl. 9, pp. 129-139.

Elliot, E. (1958) 'A ripple effect in the audiogram' Nature, 181, p. 1076.

Fukazawa, T. (1992) 'Evoked otoacoustic emissions in a nonlinear model of the cochlea' Hear. Res. 59, pp. 17-24.

Fukazawa, T. and Tanaka, Y. (1996) 'Spontaneous otoacoustic emissions in an active feed-forward model of the cochlea' Hear. Res. 95, pp. 135-143.

Furst, M. and Lapid, M. (1988) 'A cochlear model for otoacoustic emissions' J. Acoust. Soc. Am. 84, pp. 222-229.

Geisler, C. D. (1991) 'A cochlear model using feedback from motile outer hair cells' Hear. Res. 54, pp. 105-117.

Giraud, A. L., Perrin, E., Chéry-Croze, S., Chays, A. and Collet, L. (1996) 'Contralateral acoustic stimulation induces a phase advance in evoked otoacoustic emissions in humans' Hear. Res. 94, pp. 54-62.

Gold, T. (1948) 'Hearing II. The physical basis for the action of the cochlea' Proc. R. Soc. London B 135, pp. 492-498.

Guelke, R. W. and Bunn, A. E. (1985) 'A mechanism for stimulated acoustic emissions in the cochlea' Hear. Res. 19, pp. 185-189.

Gutowski, P. R., Robinson, E. A. and Treitel, S. (1978) 'Spectral estimation: Fact or fiction' in D. G. Childers (Ed.), *Modern Spectrum Analysis*, IEEE Press Selected Reprint Series, John Wiley and Sons, New York, pp. 49-53.

He, N. and Schmiedt, R. A. (1996) 'Effects of aging on the fine structure of the 2f1-f2 acoustic distortion product' J. Acoust. Soc. Am. 99, pp. 1002-1015.

He, N. and Schmiedt, R. A. (1997) 'Fine structure of the 2f1-f2 acoustic distortion product: Effects of primary level and frequency ratios' J. Acoust. Soc. Am. 101, pp. 3554-3565.

Hsu, H. P. (1995) *Signals and Systems*, McGraw-Hill, New York.

Jurzita, D. and Hemmert, W. (1992) 'Quantitative measurements of simultaneous evoked otoacoustic emissions' Acustica 77, pp. 93-99.

Kaernbach, C., König, P., Schillen, T. (1987) 'On Riccati equations describing impedance relations for forward and backward excitation in the one-dimensional cochlea model' J. Acoust. Soc. Am. 81, pp. 408-411 .

Kanis, L. J. and de Boer, E. (1993a) 'The emperor's new clothes: DP emissions in a locally-active nonlinear model of the cochlea' in H. Duifhuis, J. W. Horst, P. van Dijk, S. M. van Netten (Ed.), *Biophysics of Hair-Cell Sensory Systems*, World Scientific, Singapore, pp. 304-311.

Kanis, L. J. and de Boer, E. (1993b) 'Self-suppression in a locally active non-linear model of the cochlea: A quasilinear approach' J. Acoust. Soc. Am. 94, pp. 3199-3206.

Kanis, L. J. and de Boer, E. (1994) 'Two-tone suppression in a locally active nonlinear model of the cochlea' J. Acoust. Soc. Am. 96, pp. 2156-2165.

Kanis, L. J. and de Boer, E. (1996) 'Comparing frequency domain with time domain solutions for a locally active nonlinear model of the cochlea' J. Acoust. Soc. Am. 100, pp. 2542-2546.

Kanis, L. J. and de Boer, E. (1997) 'Frequency dependence of acoustic distortion products in a locally active model of the cochlea' J. Acoust. Soc. Am. 101, pp. 1527-1531.

Kapadia, S. and Lutman, M. E. (1999) 'Reduced audiogram 'ripple' in normally-hearing subjects with weak otoacoustic emissions' Audilogy 38, pp. 257-261.

Karlsson, K. K., Berninger, E. and Alván, G. (1991) 'The effect of quinine on pyschoacoustical tuning curves, stapedius reflexes and evoked otoacoustic emissions in healthy volunteers' Scand. Audiol. 20, pp. 83-90.

Keefe, D. H. (1997) 'Otoreflectance of the cochlea and middle ear', J. Acoust. Soc. Am. 102, pp. 2849-2859.

Kemp, D. T. (1978) 'Stimulated acoustic emissions from within the human auditory system' J. Acoust. Soc. Am. 64, pp. 1386-1391.

Kemp, D. T. (1979) 'The evoked cochlear mechanical response and the auditory microstructure - evidence for a new element in cochlear mechanics' in M. Hoke and E. de Boer (Ed.), *Models of the auditory system and related signal processing techniques*, Scand. Audiol. Suppl. 9, pp. 35-47.

Kemp, D. T. (1980) 'Toward a model for the origin of cochlear echoes' Hear. Res. 2, pp. 533-548.

Kemp, D. T. (1986) 'Otoacoustic emissions, travelling waves and cochlear mechanisms' Hear. Res. 22, pp 95-104.

Kemp, D. T., Brass, D. N. and Souter, M (1990) 'Observations on simultaneous SFOAE and DPOAE generation and suppression' in P. Dallos, C. D. Geisler, J. W. Matthew, M. A. Ruggero, C. R. Steele (Ed.), *The Mechanics and Biophysics of Hearing*, Springer-Verlag, Berlin, pp. 202-209.

Kemp, D. T. and Brown, A. M. (1983) 'A comparison of mechanical nonlinearities in the cochleae of man and gerbil from ear canal measurements' in R. Klinke and R. Hartmann (Ed.), *Hearing - Physiological Bases and Psychophysics*, Springer, Berlin, pp. 82-88.

Kemp, D. T. and Chum, R. A. (1980a) 'Observations on the generation mechanism of stimulus frequency acoustic emissions - two tone suppression.' in G. van den Brink, F. A. Bilsen (Ed.), *Psychophysical, Physiological and Behavioural Studies in Hearing*, Proceedings of the 5th International Symposium on Hearing, Delft University Press, The Netherlands, pp. 34-42.

Kemp, D. T. and Chum, R. A. (1980b) 'Properties of the generator of stimulated acoustic emissions' Hear. Res. 2, pp. 213-232.

Kolston, P. J. (1988) 'Sharp mechanical tuning in a cochlear model without negative damping' J. Acoust. Soc. Am. 83, pp. 1481-1487.

Kolston, P. J. and Smoorenburg., G. F. (1990) 'Does the cochlear amplifier produce reactive or resistive forces?' in P. Dallos, C. D. Geisler, J. W. Matthew, M. A. Ruggero, C. R. Steele (Ed.), *The Mechanics and Biophysics of Hearing*, Springer-Verlag, Berlin, pp. 96-105.

Koshigoe, S. and Tubis, A. (1983) 'Frequency-domain investigations of cochlear stability in the presence of active elements' J. Acoust. Soc. Am. 73, pp. 1244-1248.

Kringlebotn, M. (1988) 'Network model for the human middle ear' Scand. Audiol. 17, pp. 75-85.

Leeuw, A. R. and Dreschler, W. A. (1998) 'The relation between otoacoustic emissions and the broadening of the auditory filter for higher levels' Hear. Res. 126, pp. 1-10.

Lighthill, M. J. (1981) 'Energy flow in the cochlea' J. Fluid Mech. 106, pp. 149-213.

Lind, O. (1994) 'Contralateral suppression of TEOAE. Attempts to find a latency' Br. J. Audiol. 28, pp. 219-225.

Long, G. R. and Tubis, A. (1988) 'Modification of spontaneous and evoked otoacoustic emissions and associated psychoacoustic microstructure by aspirin consumption' J. Acoust. Soc. Am. 84, pp. 1343-1353.

Lonsbury-Martin, B. L., Harris, F. P., Stagner, B. B., Hawkins, D. H. and Martin, G. K. (1990) 'Distortion product emissions in humans II. Relations to acoustic immitance and stimulus frequency and spontaneous otoacoustic emissions in normally hearing subjects' Ann. Otol. Rhinol. Laryngol. Suppl. 147, pp. 15-29.

Lutman, M. E. and Deeks, J. (1999) 'Correspondence amongst microstructure patterns observed in otoacoustic emissions and Bekesy' Audilogy 38, pp. 263-266.

Manley, G. A. (1983) 'Frequency spacing of acoustic emissions: A possible explanation' in R. A. Webster and L. M. Aitken (Ed.), *Mechanisms of Hearing*, Monash U.P., Clayton, Australia, pp. 36-39.

Micheyl, C. and Collet, C. (1994) 'Interrelations between Pyschoacoustical Tuning Curves and Spontaneous and Evoked Otoacoustic Emissions' Scand. Audiol. 23, pp. 171-178.

Molenaar, D. G. , Shaw G. and Eggermont, J. J. (2000) 'Noise suppression of transient-evoked otoacoustic emissions. I. A comparison with the non-linear method' Hear. Res. 143, pp. 197-207.

Neely, S. T. (1993) 'A model of cochlear mechanics with outer hair cell motility' J. Acoust. Soc. Am. 94, pp. 137-146.

Neely, S. T. and Kim, D. O. (1986) 'A model for active elements in cochlear biomechanics' J. Acoust. Soc. Am. 79, pp. 1472-1480.

Neely, S. T., Norton, S. J., Gorga, M. P. and Jesteadt, W. (1988) 'Latency of auditory brain-stem responses and otoacoustic emissions using tone-burst stimuli' J. Acoust. Soc. Am. 82, pp. 652-656.

Newland, D. E. (1984) *An introduction to random vibrations and spectral analysis*, Longman, England.

Norton, S. J. and Neely, S. T. (1987) 'Tone-burst-evoked otoacoustic emissions from normal-hearing subjects' J. Acoust. Soc. Am. 81, pp. 1860-1872.

Patuzzi, R. (1996) 'Cochlear Micromechanics and Macromechanics' in P. Dallos, A. N. Popper, R. R. Fay (Ed.), *The Cochlea*, Springer-Verlag, pp. 186-257.

Peisl, W. (1988) 'Simulation von zeitverzögerten evozierten oto-akustischen Emissionen mit hilfe eines digitalen Innenohrmodells.' In: *Fortschritte der Akustik* - DAGA'88, DPG-GmbH, Bad Honnef, pp. 553-556.

Pickles, J. O. (1988) *Introduction to the Physiology of Hearing*, Second Edition, Academic Press Limited, London.

Probst, R., Coats A. C., Martin, G. K. and Lonsbury Martin, B. L. (1986) 'Spontaneous, click-, and toneburst-evoked otoacoustic emissions from normal ears' Hear. Res. 21, pp. 261-275.

Probst, R., Lonsbury Martin, B. L. and Martin, G. K. (1991) 'A review of otoacoustic emissions' J. Acoust. Soc. Am. 89, pp. 2027-2067.

Randall, R. B. (1987) *Frequency analysis by R. B. Randall*, Second Edition, Brüel and Kjaer, Naerum, Denmark.

Rhode, W. S. (1971) 'Observations of the vibration of the basilar membrane in squirrel monkeys using the Mössbauer technique' J. Acoust. Soc. Am. 49, pp. 1218-1231.

Rhode, W. S. and Cooper, N. P. (1993) 'Two-tone suppression and distortion production on the basilar membrane in the hook region of cat and guinea pig cochleae' Hear. Res. 66, pp. 31-45.

Rutten, W. L. C. (1980) 'Evoked acoustic emissions from within normal and abnormal human ears: Comparison with audiometric and electrocochleographic findings' Hear. Res. 2, pp. 263-271.

Schloth, E. (1983) 'Relation between spectral composition of spontaneous otoacoustic emissions and fine-structure of threshold in quiet' Acustica 53, pp. 252-256.

Sellick, P. M., Patuzzi, R. and Johnstone, B. M. (1982) 'Measurement of basilar membrane motion in the guinea pig using the Mössbauer technique' J. Acoust. Soc. Am. 72, 131-141.

Shera, C. A. and Guinan, Jr, J. J. (1999) 'Evoked otoacoustic emissions arise by two fundamentally different mechanisms: A taxonomy for mammalian OAEs' J. Acoust. Soc. Am. 105, pp. 782-798.

Shera, C. A. and Zweig, G. (1991a) 'A symmetry suppresses the cochlea catastrophe' J. Acoust. Soc. Am. 89, pp. 1276-1289.

Shera, C. A. and Zweig, G. (1991b) 'Reflection of retrograde waves within the cochlea and at the stapes' J. Acoust. Soc. Am. 89, pp. 1290-1305.

Shera, C. A. and Zweig, G. (1992b) 'An empirical bound on the compressibility of the cochlea.' J. Acoust. Soc. Am. 92, pp. 1382-1388.

Shera, C. A. and Zweig, G. (1993a) 'Noninvasive measurement of the cochlear traveling-wave ratio' J. Acoust. Soc. Am. 93, pp. 3333-3352.

Shera, C. A. and Zweig, G. (1993b) 'Order from chaos: resolving the paradox of periodicity in evoked otoacoustic emissions' in H. Duifhuis, J. W. Horst, P. van Dijk, S. M. van Netten (Ed.), *Biophysics of Hair-Cell Sensory Systems*, World Scientific, Singapore, pp. 54-63.

Strube, H. W. (1985) 'A computationally efficient basilar-membrane model' Acustica 58, pp. 207-214.

Strube, H. W. (1989) 'Evoked otoacoustic emissions as cochlear Bragg reflections' Hear. Res. 38, pp. 35-46.

Sutton G. J. (1985) 'Suppression effects in the spectrum of evoked oto-acoustic emissions' Acustica 58, pp. 57-63.

Talmadge, C. L. and Tubis, A. (1993) 'On modelling the connection between spontaneous and evoked otoacoustic emissions' in H. Duifhuis, J. W. Horst, P. van Dijk, S. M. van Netten (Ed.), *Biophysics of Hair-Cell Sensory Systems*, World Scientific, Singapore, pp. 25-32.

Talmadge, C. L., Tubis A., Long, G. R. and Piskorski, P. (1998) 'Modeling otoacoustic emission and hearing threshold fine structures' J. Acoust. Soc. Am. 104, pp. 1517-1543.

Talmadge, C. L., Tubis A., Long, G. R. and Tong, C. (2000) 'Modeling the combined effects of basilar membrane nonlinearity and roughness on stimulus frequency otoacoustic emission fine structures' J. Acoust. Soc. Am. 108, pp. 2911-2932.

Tavartkiladze, G. A., Frolenkov, G. I., Kruglov, A. V. and Artamasov, S. V. (1994) 'Ipsilateral suppression effects of transiently evoked otoacoustic emission' Br. J. Audiol. 28, pp. 193-204.

Thomson, W. T. (1981) *Theory of vibration with applications*, Second Edition, Prentice-Hall, Englewood Cliffs, New Jersey.

van Hengel, P. W. J., Duifhuis, H. and van den Raadt, M. P. M. G. (1996) 'Spatial periodicity in the cochlea: The result of interaction of spontaneous emissions?' J. Acoust. Soc. Am. 99, pp. 3566-3571.

van Hengel, P. W. J. and Maat, A. (1993) 'Periodicity in frequency spectra of click evoked and spontaneous OAE's; theory meets experiment' in H. Duifhuis, J. W. Horst, P. van Dijk, S. M. van Netten (Ed.), *Biophysics of Hair-Cell Sensory Systems*, World Scientific, Singapore, pp. 47-53.

Viergever, M. A. and de Boer, E. (1987) 'Matching impedance of a nonuniform transmission line: Application to cochlear modeling' (letters) J. Acoust. Soc. Am. 81, pp. 184-186.

Viergever, M. A. and Diependaal, R. J. (1986) 'Quantitative validation of cochlear models using the Liouville-Green approximation' Hear. Res. 21, pp. 1-15.

Wada, H., Nakajima, T. and Ohyama, K. (1999) 'Effect of irregularities in basilar membrane impedance on TEOAEs: Theoretical Considerations' ORL. J. Otorhinolaryngol. Relat. Spec. 61, pp. 252-258.

Wilson, J. P. (1980a) 'Evidence for a cochlear origin for acoustic re-emissions, threshold fine-structure and tonal tinnitus' Hear. Res. 2, pp. 233-252.

Wilson, J. P. (1980b) 'Model for cochlear echos and tinnitus based on an observed electrical correlate' Hear. Res. 2, pp. 527-532.

Wit, H. P., van Dijk, P. and Avan, P. (1994) 'On the shape of (evoked) otoacoustic emission spectra' Hear. Res. 81, pp. 208-214.

Zweig, G., (1990) 'The Impedance of the Organ of Corti' in P. Dallos, C. D. Geisler, J. W. Matthew, M. A. Ruggero, C. R. Steele (Ed.), *The Mechanics and Biophysics of Hearing*, Springer-Verlag, Berlin, pp. 362-369.

Zweig, G. (1991) 'Finding the Impedance of the Organ of Corti' J. Acoust. Soc. Am. 89; pp. 1229-1254.

Zweig, G., Lipes, R. and Pierce, J. R. (1976) 'The cochlear compromise' J. Acoust. Soc. Am. 59; pp. 975-982.

Zweig, G. and Shera, C. A. (1995) 'The origin of periodicity in the spectrum of evoked otoacoustic emissions' J. Acoust. Soc. Am. 98, pp. 2018-2047.

Zwicker, E. (1979) 'A model describing nonlinearities in hearing by active processes with saturation at 40 dB' Biol. Cybern. 35, pp. 243-250.

Zwicker, E. (1986) 'Otoacoustic emissions in a nonlinear cochlear hardware model with feedback' J. Acoust. Soc. Am. 80, pp. 154-162.

Zwicker, E. (1988) 'Otoacoustic emissions and cochlear travelling waves' in J. P. Wilson and D. T. Kemp (Ed.), *Cochlear mechanisms - structure, function, and models*, (NATO ASI Series), Plenum Press, New York, pp. 359-366.

Zwicker, E. (1990) 'On the frequency separation of simulataneously evoked otoacoustic emissions' consecutive extrema and its relation to cochlear travelling waves' J. Acoust. Soc. Am. (letter), 88, pp. 1639-1641.

Zwicker, E. and Lumer, G. (1985) 'Evaluating traveling wave characteristics in man by an active nonlinear cochlea preprocessing model' in J. B. Allen, J. L. Hall, A. Hubbard,

S. T. Neely and A. Tubis (Ed.), *Peripheral Auditory Mechanisms* (Lecture Notes in Biomathematics, 64), Springer, Berlin, pp. 250-257.

Zwicker, E. and Peisl, W. (1990) 'Cochlear preprocessing in analog models, in digital models and in human inner ear' Hear. Res. 44, pp. 209-216.

Zwicker, E. and Schloth, E. (1984) 'Interrelation of different otoacoustic emissions' J. Acoust. Soc. Am. 75, pp. 1148-1154.

Zwicker, E. and Wesel, J. (1990) 'The effect of "addition" in suppression of delayed evoked oto-acoustic emissions and in masking' Acustica 70, pp. 189-196.

Appendix I Derivation of the Longwave Equation

Derivations of the longwave equation can be found in Lighthill (1981) and de Boer (1996). In order to indicate the precise variant of the longwave equation that has been adopted in this these, a full derivation is presented here. The notation introduced in sections 4.3 and 4.4 is used throughout this section.

I-1 Derivation of the Wave Equation for Antisymmetric Loading

Consider first the cochlear fluid in the scala vestibuli under antisymmetric (push-pull) loading. In this case, from [4.1] with $p_s(x,y,t) \equiv 0$, the fluid pressure in the scala vestibuli is simply equal to the semi-difference pressure, $p_d(x,y,t)$. The fluid then obeys the equation of conservation of mass:

$$\frac{\partial u(x,y,t)}{\partial x} + \frac{\partial v(x,y,t)}{\partial y} = 0 \qquad \text{(conservation of mass)} \qquad \text{[I-1]}$$

and the equations of conservation of momentum in two directions:

$$\frac{\partial p_d(x,y,t)}{\partial x} = -\rho \frac{\partial u(x,y,t)}{\partial t} \qquad \text{(conservation of } x\text{-momentum)} \qquad \text{[I-2]}$$

$$\frac{\partial p_d(x,y,t)}{\partial y} = -\rho \frac{\partial v(x,y,t)}{\partial t} \qquad \text{(conservation of } y\text{-momentum)} \qquad \text{[I-3]}$$

together with the boundary conditions at the stapes and helicotrema:

$$\begin{aligned} u(x,y,t)\big|_{x=0} &= u_{St}(t) \qquad \forall y,t \\ u(x,y,t)\big|_{x=L} &= 0 \qquad\quad \forall y,t \end{aligned} \qquad \text{[I-4]}$$

and at the CP and ceiling of the scala vestibuli:

$$\begin{aligned} v(x,y,t)\big|_{y=0} &= v_{CP}(x,t) \qquad \forall x,t \\ v(x,y,t)\big|_{y=H} &= 0 \qquad\quad \forall x,t \end{aligned} \qquad \text{[I-5]}$$

where it has been assumed that the stapes x-velocity is independent of y. It has also been assumed that the fluid is incompressible (such that ρ is constant) and that all fluid displacements and velocities are small, such that nonlinear terms in the momentum equations are neglected.

In the longwave approximation, it is assumed that the y-velocity, $v(x,y,t)$, varies linearly across the scala. It then follows from [I-5] that:

$$v(x,y,t) = (1 - \frac{y}{H}) v_{CP}(x,t) \qquad \text{[I-6]}$$

It can be shown that in adopting this assumption, the conservation of y-momentum is violated (i.e., that [I-3] is *not* satisfied). This is approximately valid provided that $v(x,y,t) << u(x,y,t)$, which occurs when H is much smaller than the wavelength of the TW. By differentiating [I-6] w.r.t. y and substituting into [I-1] it follows that:

$$\frac{\partial u(x,y,t)}{\partial x} = \frac{v_{CP}(x,t)}{H} \qquad \text{[I-7]}$$

Since the right-hand side of [I-7] is independent of y, so too is the left-hand side. It follows from [I-7] evaluated at the stapes and from the assumption that stapes velocity is independent of y that $u(x,y,t)$ is independent of y. Also, from [I-2], the pressure also becomes independent of y. Thus, by adopting the longwave assumption, it follows that both fluid pressure and x-velocity are uniform across the scala. Henceforth the y-independent variable will be dropped from the equations.

Differentiating [I-7] w.r.t. t, and [I-2] w.r.t. x, we obtain:

$$\begin{aligned}\frac{\partial^2 p_d(x,t)}{\partial x^2} &= -\rho \frac{\partial^2 u(x,t)}{\partial x \partial t} \\ &= -\frac{\rho}{H}\frac{\partial v_{CP}(x,t)}{\partial t}\end{aligned} \qquad \text{[I-8]}$$

Transforming [I-8] into the frequency domain gives:

$$\frac{d^2 \boldsymbol{p_d}(x,\omega)}{dx^2} = -i\frac{\omega\rho}{H}\boldsymbol{v}_{CP}(x,\omega) \qquad \text{[I-9]}$$

But $\boldsymbol{p_d}$ equals half the pressure difference across the CP, which is related to the CP velocity by the impedance relation:

$$\boldsymbol{v_{CP}}(x,\omega) = -2\frac{\boldsymbol{p_d}(x,\omega)}{\boldsymbol{Z_{CP}}(x,\omega)} \qquad \text{[I-10]}$$

where the –2 in [I-10] arises from the sign convention and from the definition of the semi-difference pressure. Note that the boundary condition at the helicotrema can be represented by specifying an appropriate value of $\boldsymbol{Z_{CP}}(x,\omega)$ at $x = L$. Then substituting [I-10] into [I-9] then gives the wave equation.

$$\frac{d^2 \boldsymbol{p_d}(x,\omega)}{dx^2} = i\omega\rho\frac{2\boldsymbol{p_d}(x,\omega)}{\boldsymbol{Z}_{CP}(x,\omega)} \qquad \text{[I-11]}$$

Rearranging [I-11] gives the form introduced in [4.4]:

$$\frac{d^2 \boldsymbol{p}_d(x,\omega)}{dx^2} + \boldsymbol{k}^2(x,\omega)\boldsymbol{p}_d(x,\omega) = 0$$

$$\boldsymbol{k}(x,\omega) \equiv \frac{\omega}{\boldsymbol{c}_{TW}(x,\omega)} \qquad \text{[I-12]}$$

$$\boldsymbol{c}_{TW}^2(x,\omega) \equiv \frac{i\omega \ HZ_{CP}(x,\omega)}{2\rho}$$

Solving [I-12] together with the remaining boundary conditions [I-4] then yields $\boldsymbol{p}_d(x,\omega)$, $\boldsymbol{u}(x,\omega)$, $v_{CP}(x,\omega)$ and $v(x,y,\omega)$.

I-2 The Effect of Symmetric Loading

In the above derivation it was first assumed that the loading was antisymmetric. It will now be shown that the situation remains unchanged when introducing a symmetric loading component, since this alters neither the fluid velocities nor the pressure gradient.

For a purely symmetric loading, with $p_d(x,y,t) \equiv 0$ in [4.1], it follows that the fluid pressure in the scala vestibuli is simply equal to the semi-sum pressure, $p_s(x,y,t)$. The x- and y-momentum equations become:

$$\frac{\partial p_s(x,y,t)}{\partial x} = -\rho \frac{\partial u(x,y,t)}{\partial t} \qquad \text{(conservation of } x\text{-momentum)} \qquad \text{[I-13]}$$

$$\frac{\partial p_s(x,y,t)}{\partial y} = -\rho \frac{\partial v(x,y,t)}{\partial t} \qquad \text{(conservation of } y\text{-momentum)} \qquad \text{[I-14]}$$

However, from symmetry in the scala vestibuli and scala tympani and from the assumption of fluid incompressibility, it follows that $u(x,y,t) = 0$ and $v(x,y,t) = 0$ for all x,y and t. Therefore:

$$\frac{\partial p_s(x,y,t)}{\partial x} = 0 \qquad \text{(conservation of } x\text{-momentum)} \qquad \text{[I-15]}$$

$$\frac{\partial p_s(x,y,t)}{\partial y} = 0 \qquad \text{(conservation of } y\text{-momentum)} \qquad \text{[I-16]}$$

And therefore p_s is independent of both x and y.

I-3 The Complete Solution to Asymmetric Loading

For a general loading, as specified by a velocity at the stapes and an impedance at the round window, the complete solution for the fluid pressure is as follows.

$$\begin{aligned} \boldsymbol{p}_{RW}(\omega) &\equiv \boldsymbol{p}(x,\omega)\,|_{x=0;y<H} = 0 \text{ for all } \omega \\ \boldsymbol{p}_{St}(\omega) &\equiv \boldsymbol{p}(x,\omega)\,|_{x=0;y>H} \end{aligned} \qquad \text{[I-17]}$$

and, by the definition in [4.1], these pressures can be split into the semi-sum and semi-difference pressures:

$$\begin{aligned} \boldsymbol{p}_{St}(\omega) &\equiv \boldsymbol{p}_s(\omega) + \boldsymbol{p}_d(x,\omega)\,|_{x=0} \\ \boldsymbol{p}_{RW}(\omega) &\equiv \boldsymbol{p}_s(\omega) - \boldsymbol{p}_d(x,\omega)\,|_{x=0} \end{aligned} \qquad \text{[I-18]}$$

where the semi-sum pressure is independent of x, as shown above. The boundary conditions at the stapes and round window become:

$$\begin{aligned} \boldsymbol{p}_{RW}(\omega) &\equiv -\boldsymbol{Z}_{RW}(\omega)\boldsymbol{u}(x,\omega)\,|_{x=0;y<H} = 0 \text{ for all } \omega \\ &= -\boldsymbol{Z}_{RW}(\omega)\boldsymbol{u}_{St}(\omega) \end{aligned} \qquad \text{[I-19]}$$

$$\begin{aligned} \boldsymbol{u}_{St}(\omega) &= \frac{1}{-i\rho\omega}\left.\frac{d\boldsymbol{p}(x,\omega)}{dx}\right|_{x=0;y>H} \\ &= \frac{1}{-i\rho\omega}\left.\frac{d\boldsymbol{p}_d(x,\omega)}{dx}\right|_{x=0} \end{aligned} \qquad \text{[I-20]}$$

where the minus sign in [I-19] arises from the antisymmetry of the stapes and round window velocities, which applies even in the general loading case, due to fluid incompressibility. Equation [I-20] is the frequency domain version of the x-momentum equation, [I-2]. It follows from [I-20] that $\boldsymbol{p}_d$ is dependent on $\boldsymbol{u}_{St}$ only and not on $\boldsymbol{Z}_{RW}$. It can be solved from this boundary condition and the wave equation [I-12]. Having solved for $\boldsymbol{p}_d$, the semi-sum pressure then follows from [I-18] and [I-19] and [I-20]:

$$\begin{aligned} \boldsymbol{p}_s(\omega) &= \boldsymbol{p}_{RW}(\omega) + \boldsymbol{p}_d(x,\omega)\,|_{x=0} \\ &= -\boldsymbol{Z}_{RW}(\omega)\boldsymbol{u}_{St}(\omega) + \boldsymbol{p}_d(x,\omega)\,|_{x=0} \\ &= \frac{\boldsymbol{Z}_{RW}(\omega)}{i\rho\omega}\frac{d\boldsymbol{p}_d(x,\omega)}{dx} + \boldsymbol{p}_d(x,\omega)\,|_{x=0} \end{aligned} \qquad \text{[I-21]}$$

Thus, in the case where $\boldsymbol{Z}_{RW}$ is negligible, the semi-sum pressure is simply equal to the semi-difference pressure at the stapes.

Appendix II Listings of Matlab Programs for Cochlear Models

Two listing are given here. The first is for a linear cochlear model. The second is for the nonlinear model for two-tone suppression. Self-suppression can be simulated using the two-tone suppression model, with the suppressor tone set to zero.

```
%
% _____ Linear Cochlear Model _____
% IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
%
% Matlab Version 4.2
%
% Based on:
%    Neely & Kim (1986)
%    Kanis & de Boer (1994)
%    Zweig (1991)
%    Talmadge et al. (1998)
%    Kringlebotn (1988)
%
% B.Lineton 5-3-1999
%
%  returns the ear canal pressure and volume velocity
%
%  pEC(ireal,ifreq); QEC(ireal,ifreq)
%
%  where ireal is the index of the realization
%        ifreq is the index of stimulus frequency
%
%
%  also returns (for a single realization):
%
%  pED(ifreq) QED(ifreq)   : ear drum pressure & vol. vel.
%  pSt(ifreq) QSt(ifreq)   : stapes pressure & vol. vel.
%  vCP(ix,ifreq)           : CP velocity              at place & freq
%  u(ix,ifreq)             : fluid x-velocity         at place & freq
%  p(ix,ifreq)             : fluid semi-diff pressure at place & freq
%
%  ======================== User Defined Inputs ========================
matname='thesis\test_01.mat';  % Output mat filename
freq=[1500:4:3000];       % Stimulus frequencies [Hz]
CAFlag=1;                 % Coch Amp Flag {1=K & de B; 2=Talmadge}
CAFactor=1.0;             % Coch Amp Gain Factor (0=>Passive)
Nx=1024;                  % Number of points on CP
L=30e-3;                  % Length of cochlea [m]
ASt=1.e-6;                % Cross sectional area of stapes [m^2]
rho=1000;                 % Density of Fluid [kg/m^3]
if CAFlag==1              % ==> KdB-1994 & NK-1986 Coch. Amp.
  W=1.e-3;                % Width  of cochlea [m]
  H=1.e-3;                % Height of cochlea [m]
  kom=150.0;              % Place-frequency length parameter [m^-1]
  delta=0.4;              % Damping ratio=2.Zeta=R/sqrt(S.M)
  m0=0.5;                 % Areal density of BM [kg/m^2]
  s0=1.E10;               % stiffness/unit area of BM [kg/m^2/s^2]
  mc0=0.06;               % c0 parameter [kg/m^2]
  deltaSC=0.14;           % Stereocilia damping
  sigma=0.7;              % OHC Resonance shift ratio
  e0=4.28e-5;             % Active OHC impedance parameter [kg.m^-2]
  d0=1404;                % Active OHC impedance parameter [kg.s^-1]
elseif CAFlag==2;         % ==> T-1998 & Z-1991 Coch Amp.
  kom=138.2;              % Place-frequency length parameter [m^-1]
  kgamma=kom;             % Place-damping length parameter [m^-1]
  gamma0=5035;            % Damping Parameter at x=0 (r0/m0), [s^-1]
  gamma1=100;             % Non-scaling Damping Param. (r0/m0), [s^-1]
  omegaC0=20.8e3*2*pi;    % Characteristic frequency at stapes [rad/s]
  omegaC1=-145.5*2*pi;    % Non-scaling char frequency parameter [rad/s]
```

```
   rhofast=0.16;          % Fast coefficient
   phifast=0.24*2*pi;     % Fast angle [rad]
   rhoslow=0.1416;        % Slow coefficient
   phislow=1.742*2*pi;    % Slow angle [rad]
   m0=0.055;              % Areal density of BM [kg/m^2]
   W=0.029e-2;            % Avg width of BM [m]
   H=1.1E-6/0.029e-2;     % Height of cochlea [m]
end
ScatFlg=1;                % Scattering flag {0= no inhomos; 1= Z&SI;
                          % 2=StubeI; 3=StrubeII; 4=Point; 5=Z&SII}
ScatSize=0.01;            % Standard dev as a fraction of nominal
ScatLen=10.e-6;           % Length Parameter [m] (meaning depends on
                          % ScatFlag)
% ScatLen=3.46e-4;        % Strube spatial period
% ScatLen=0.014;          % Location of Point Reflection Site [m]
% ScatLenBW=4*ScatLen;    % For StrubeII only: Spatial Bandwidth [m]
Navg=1;                   % Number of realizations in the ensemble
AEC=0.4*0.01^2;           % C.S.A. ear canal [m^2]
GLME=1/1.4;               % Lever ratio of ossicular chain
MidEarFlg=0;              % Middle Ear flag {0=Transparent, 3=Kringlebotn}
rho0=1.225;               % Density of air [kg/m^3]
c0=340;                   % Speed of sound in air [m/s]
LEC=0.0001;               % Length of ear canal [m]
QSrc0=4.8019e-008;        % Source Short Circuit Volume Velocity [m^3/s]
YSrc0=0;                  % Source Admittance [m^3/s/Pa]

% =================== End of User Inputs ============================
%
%                    Derived & Preset Quantities
%                    IIIIIIIIIIIIIIIIIIIIIIIII
%
Nfreq=length(freq);              % Number of frequencies

WH=W*H;                          % csa of cochlea
xs=L/(Nx-1);                     % Length step [m]
x=[0:Nx-1]*xs;                   % All NX points along the cochlea
x2=x(1:Nx-1);                    % NX-1 points exluding h/trema
len=1/kom;                       % place-freq length [m]
if CAFlag==1;                    % ==> KdB-1994 & NK-1986 Coch Amp.
  omegaC0=sqrt(s0/m0);           % Char frequency at stapes [rad/s]
  omegaC=omegaC0*exp(-kom*x2);   % Char freq along BM [rad/s]
  r0=delta*sqrt(m0*s0);          % Damping/unit area at stapes [Ns/m^3]
  m1=m0;                         % Mass/unit area along CP [kg/m^2]
  s1=m1*omegaC.^2;               % Stiffness/unit area along CP [N/m^3]
  r1=delta*sqrt(m1*s1);          % Damping/unit area along CP [Ns/m^3]
elseif CAFlag==2;                % ==> T-1998 & Z-1991 Coch Amp.
  omegaC=omegaC0*exp(-kom*x2)+omegaC1;
  gamma=gamma0*exp(-kom*x2)+gamma1;
  s0=m0*omegaC(1)^2;             % Stiffness/unit area at stapes [N/m^3]
  r0=m0*gamma(1);                % Damping/unit area at stapes [Ns/m^3]
  m1=m0;                         % Mass/unit area along CP [kg/m^2]
  s1=m1*omegaC.^2;               % Stiffness/unit area along CP [N/m^3]
  r1=gamma*m1;                   % Damping/unit area along CP [Ns/m^3]
  mkappaf=rhofast*s1;            % Fast feedback stiff'ss along CP [N/m^3]
  tauf=phifast./omegaC;          % Fast feedback delay [s]
  mkappas=rhoslow*s1;            % Slow feedback stiff'ss along CP [N/m^3]
  taus=phislow./omegaC;          % Slow feedback delay [s]
end;
QSrc=QSrc0*ones(size(freq));
YSrc=YSrc0*ones(size(freq));
ZHel=1e-12;  % Effective impedance at helicotrema (make v.small)
%
% === Two Port Representation of Ear Canal and Middle Ear ===
%
GOME=AEC/(ASt*GLME);  % Overall middle ear static pressure gain
if MidEarFlg==0  % Transparent
  TEDoSt_11=ones(size(freq)) *1/GOME;  % <=== Stapes to Ear Drum
Transmission Coeff.
  TEDoSt_12=zeros(size(freq));
  TEDoSt_21=zeros(size(freq));
```

```
  TEDoSt_22=ones(size(freq)) *GOME;
elseif MidEarFlg==3  % Kringlebotn
  % NB: inputs to t_kring are in CGS units, outputs are in SI units
  milli=1.e-3;  micro=1.e-6;
  La=1.00*milli;  Ca=3.90*micro;   Ra=60;
                  Ct=0.40*micro;
  Ld=7.50*milli;
  Ls=66.0*milli;  Cs=0.30*micro;   Rs=20;
                  Cr=1.30*micro;   Rr=120;
                  Cm=0.38*micro;   Rm=120;
  Lo=22.0*milli;                   Ro=20;
                  Ci=0.30*micro;   Ri=6000;
  Lc=46.0*milli;  Cc=0.56*micro;   Rc=330;

MidEarParam=[ASt,AEC,GLME,La,Ca,Ra,Ct,Ld,Ls,Cs,Rs,Cr,Rr,Cm,Rm,Lo,Ro,Ci,Ri
,Lc,Cc,Rc];
  [TEDoSt_11,TEDoSt_12,TEDoSt_21,TEDoSt_22]=t_kring(MidEarParam,freq);
end;
% === Ear Canal ===
rhoc0=rho0*c0;
kL=2*pi*freq*LEC/c0;
sinkL=sin(kL);
coskL=cos(kL);
tankL=tan(kL);
% === Transmission Matrices ===
TECoED_11=coskL;                  % <=== Ear Drum
TECoED_12=i*sinkL*rhoc0/AEC;      %          to
TECoED_21=i*sinkL*AEC/rhoc0;      %      Ear Canal
TECoED_22=coskL;
TECoSt_11=TECoED_11.*TEDoSt_11+TECoED_12.*TEDoSt_21; % <=== Stapes
TECoSt_12=TECoED_11.*TEDoSt_12+TECoED_12.*TEDoSt_22; %        to
TECoSt_21=TECoED_21.*TEDoSt_11+TECoED_22.*TEDoSt_21; %     Ear Canal
TECoSt_22=TECoED_21.*TEDoSt_12+TECoED_22.*TEDoSt_22;
TSrcSt_11=TECoED_11.*TEDoSt_11+TECoED_12.*TEDoSt_21; % <=== Stapes to
TSrcSt_12=TECoED_11.*TEDoSt_12+TECoED_12.*TEDoSt_22; %        to
TSrcSt_21=TECoED_21.*TEDoSt_11+TECoED_22.*TEDoSt_21; % Acoustic Source
TSrcSt_22=TECoED_21.*TEDoSt_12+TECoED_22.*TEDoSt_22;
% Refer the Acoustic source to the stapes.
QStSrc=QSrc./(TSrcSt_12.*YSrc+TSrcSt_22);                          % Q @ St
YStSrc=(TSrcSt_11.*YSrc+TSrcSt_21)./(TSrcSt_12.*YSrc+TSrcSt_22); % Y @ St
%
%              Cochlear Response Calculation
%              IIIIIIIIIIIIIIIIIIIIIIIIIII
%
pECens=zeros(Navg,Nfreq); QECens=zeros(Navg,Nfreq);
randn('seed',0);  %reset seed
for iavg=1:Navg
  if ScatFlg==1|ScatFlg==5     % Z & S broad band random scattering
    inhomo=randn(size(x2));
    Wn=2*xs/ScatLen;  % LP cut-off spatial freq /(1/2 sampling rate)
    if Wn<1.0
      [BB,AA]=butter(1,Wn);  % 1st order butterworth low pass filter
      inhomo=filtfilt(BB,AA,inhomo);
    end
  elseif ScatFlg==2 % Strube periodic corrugations
    Kappa=1/ScatLen;
    inhomo=cos(2*pi*Kappa*x2);
  elseif ScatFlg==3 % Strube narrow-band corrugations
    inhomo=randn(size(x2));
    Kappa=1/ScatLen;
    KappaBW=1/ScatLenBW;
    Wn=[Kappa-KappaBW/2,Kappa+KappaBW/2]*2*xs;
    if Wn<1.0
      [BB,AA]=butter(1,Wn);  %1st order butterworth bandpass filter
      inhomo=filtfilt(BB,AA,inhomo);
    end
  elseif ScatFlg==4 % Point reflection site
    inhomo=0.5*(1+sign(x2-ScatLen)).*hamming(length(x2)).';
  else % smooth CP
    inhomo=zeros(size(x2));
```

```
    end
    % loop over frequencies
    p_2=zeros(Nx,Nfreq);  vCP=zeros(Nx,Nfreq);  u=zeros(Nx,Nfreq);
    p=zeros(Nx,1); rhs=zeros(Nx,1);  rnew=zeros(Nx,1);  a=zeros(Nx,1);
    clc
    stepop=2;
    t0=clock;
    for ifr=1:Nfreq;
      pcent=100*ifr/Nfreq;
      if (rem(ifr,stepop)==0)    % o/p only every 2 percent (for pcentop=2)
        str1= int2str(pcent);
        str2= num2str(etime(clock,t0)/60);
        str2 = str2(1:min(5,length(str2)));
        home,disp(...
        ['Calculation is ',str1,' % complete after ',str2,' mins'])
      end
      omega=2*pi*freq(ifr);
      iomega=i*omega;
      ZPass=s1/iomega+r1+iomega*m1;  % Passive CP impedance
      if CAFlag==1;                  % Coch Amp CP impedance
          beta=omega./omegaC;
          Gxom=d0*(1+i*beta)./(deltaSC+i*(beta-sigma^2./beta));
          ZCA=-e0*omegaC.*Gxom;
      elseif CAFlag==2;
          ZCA= (mkappaf.*exp(-iomega*tauf) ...
               +mkappas.*exp(-iomega*taus))./iomega;
      end
      if ScatFlg~=5    % Scattering impedance based on damping
        ZScat=ScatSize*inhomo.*r1;
      else             % Scattering impedance based on stiffness
        ZScat=ScatSize*inhomo.*s1/iomega;
      end
      ZCP=ZPass+ZCA*CAFactor+ZScat;
      ZCP=[ZCP,ZHel];
      b=2*iomega*rho*xs^2./(H*ZCP);
      A11=-1-b(1)-2*iomega*rho*xs*YStSrc(ifr)/WH;
      A12=1;
      rhs(1)=-QStSrc(ifr)*iomega*rho*xs/WH;
      rnew(1)=rhs(1)/A11;
      a(1)=-A12/A11;
      for ix=2:Nx-1;
        a(ix)=-1/(-2-b(ix)+a(ix-1));
        rnew(ix)=(rnew(ix-1)-rhs(ix))*a(ix);
      end;
      a(Nx)=-1/(-1-b(Nx)+a(Nx-1));
      rnew(Nx)=a(Nx)*(rnew(Nx-1)-rhs(Nx));
      p(Nx)=rnew(Nx);
      for ix=Nx-1:-1:1;
        p(ix)=rnew(ix)+a(ix)*p(ix+1);          % fluid pressure
      end;
      vCP(:,ifr)=(-2*p)./ZCP.';                % CP velocity
      u(2:Nx,ifr)=diff(p)/(-iomega*rho*xs);    % fluid x-velocity
      u(1,ifr)=u(2,ifr)-vCP(1,ifr)*xs/H;
      p_2(:,ifr)=p;
    end; % end loop over freq
    clear p;p=p_2;clear p_2;
    %
    %              Ear Canal Pressure Calculation
    %              IIIIIIIIIIIIIIIIIIIIIIIIIIIII
    %
    pSt=2*p(1,:);
    QSt=u(1,:)*WH;
    pED=TEDoSt_11.*pSt+TEDoSt_12.*QSt;
    QED=TEDoSt_21.*pSt+TEDoSt_22.*QSt;
    pEC=TECoED_11.*pED+TECoED_12.*QED;
    QEC=TECoED_21.*pED+TECoED_22.*QED;
    pECens(iavg,:)=pEC;
    QECens(iavg,:)=QEC;
  end % end ensemble avg
  command=['save ',matname]; eval(command);
```

```
% IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
% ========================================================================
% IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII

function
[TEDoSt_11,TEDoSt_12,TEDoSt_21,TEDoSt_22]=t_kring(MidEarParam,freq);
%
% Kringlebotn's Middle Ear Model
% Calculate 2-port transmission matrix, TEDoSt, between stapes
% and ear drum such that:
%
% pED = [TEDoSt] pSt
% QED            QSt;

%================================ Parameters (CGS units) ===

ASt  =MidEarParam( 1);
AEC  =MidEarParam( 2);
GLME =MidEarParam( 3);
La   =MidEarParam( 4);
Ca   =MidEarParam( 5);
Ra   =MidEarParam( 6);
Ct   =MidEarParam( 7);
Ld   =MidEarParam( 8);
Ls   =MidEarParam( 9);
Cs   =MidEarParam(10);
Rs   =MidEarParam(11);
Cr   =MidEarParam(12);
Rr   =MidEarParam(13);
Cm   =MidEarParam(14);
Rm   =MidEarParam(15);
Lo   =MidEarParam(16);
Ro   =MidEarParam(17);
Ci   =MidEarParam(18);
Ri   =MidEarParam(19);
Lc   =MidEarParam(20);
Cc   =MidEarParam(21);
Rc   =MidEarParam(22);

%================================ Impedances ===

GOME=AEC/(ASt*GLME);  % Overall middle ear static pressure gain
omega=2*pi*freq;
iomega=i*omega;

Z1a=Ra+iomega*La+1./(iomega*Ca);
Z1b=             1./(iomega*Ct);
Z1c=   iomega*Ld               ;
Z1d=Rs+iomega*Ls+1./(iomega*Cs);
Z1e=Rr          +1./(iomega*Cr);
Z2 =Rm          +1./(iomega*Cm);
Z3 =Ro+iomega*Lo               ;
Z4 =Ri          +1./(iomega*Ci);
Z5a=   iomega*Lc+1./(iomega*Cc);
Z5b=Rc                         ;

Z1=Z1a.*Z1b./(Z1a+Z1b)+Z1c+Z1d.*Z1e./(Z1d+Z1e);
Z5=Z5a+Z5b;

%=================== Middle Ear Transmission matrix ===

Z_eff0=Z1+Z2.*(Z3+Z4)./(Z2+Z3+Z4);
Z_eff1=Z4.*Z5a./(Z4+Z5a);
Z_eff2=Z2.*(Z3+Z_eff1)./(Z2+Z3+Z_eff1);
Z_eff3=Z1+Z_eff2;

T11=1./((1-Z1./Z_eff0).*Z4./(Z3+Z4));
T21=1./(Z_eff0.*(1-Z1./Z_eff0).*Z4./(Z3+Z4));
```

```
T22=Z5a.*(Z3+Z_eff1)./(Z_eff1.*Z_eff2);
T12=T22.*Z_eff3;

TEDoSt_11= T11/GOME;
TEDoSt_12= T12*GOME;
TEDoSt_21= T21/GOME;
TEDoSt_22= T22*GOME;

% convert to SI units

TEDoSt_12= TEDoSt_12*1.e5;
TEDoSt_21= TEDoSt_21/1.e5;

return
```

```
% _____ Two-tone Suppression Cochlear Model _____
% IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
%
% Matlab Version 4.2
%
% Based on:
%    Neely & Kim (1986)
%    Kanis & de Boer (1993, 1994, 1996)
%    Zweig (1991)
%    Talmadge et al. (1998)
%    Kringlebotn (1988)
%
% B.Lineton 19-7-1999
%
%  returns the ear canal pressure and volume velocity for a single
%  realization
%
%  pEC(ifreq1,itone); QEC(ifreq1,itone)
%
%  where ifreq1 is the index of probe frequency
%        itone  is the index of the tone number:
%               1=> probe tone
%               2=> suppressor tone
%
%  ======================== User Defined Inputs ========================
%
% === KdB-1994 & NK-1986 Coch. Amp. ===
%
matroot='thesis\testb_03'; % mat filname root
IDstart=0;                 % file ID
Nfreq1=3;                  % number of frequencies in sweep
OctL=0;                    % lower freq in octaves
OctH=1;                    % upper freq in octaves
SPLnom=[45,30];            % desired SPL probe & suppressor in ear canal
Nx=256*4;                  % Number of points along the CP
L=30e-3;                   % Length of cochlea [m]
ASt=1.e-6;                 % Cross sectional area of stapes [m^2]
rho=1000;                  % Density of Fluid [kg/m^3]
W=1.e-3;                   % Width  of cochlea [m]
H=1.e-3;                   % Height of cochlea [m]
kom=150.0;                 % Place-frequency length parameter [m^-1]
delta=0.4;                 % Damping ratio=2.Zeta=R/sqrt(S.M)
m0=0.5;                    % Areal density of BM [kg/m^2]
s0=1.E10;                  % Stiffness/unit area of BM [kg/m^2/s^2]
mc0=0.06;                  % c0 parameter [kg/m^2]
deltaSC=0.14;              % Stereocilia damping
sigma=0.7;                 % Shift of OHC resonance wrt. BM resonance.
e0=4.28e-5;                % Active OHC impedance parameter [kg.m^-2]
d0=1404;                   % Active OHC impedance parameter [kg.s^-1]
ScatFlg=1;                 % Scattering flag {0= no inhomos; 1= ZSI;
%                          %  2=StubeI; 3=StrubeII; 4=Point; 5=Z&SII}
ScatSize=0.01;             % Standard dev as a fraction of nominal
ScatLen=10.e-6;            % Length Parameter [m]
realiz=22;                 % Realization number (for cf linear model)
AEC=0.4*0.01^2;            % C.S.A. ear canal [m^2]
GLME=1/1.4;                % Lever ratio of ossicular chain
MidEarFlg=3;               % Middle ear flag (0=Perfect transformer,
%                                             3=Kringlebotn)
rho0=1.225;                % Density of air [kg/m^3]
c0=340;                    % Speed of sound in air [m/s]
LEC=0.0001;                % Length of ear canal [m]
YSrc0=0;                   % Source Admittance [m^3/s/Pa]
% ==================== End of User Inputs ============================
freq1=1500*2.^(OctL+[0:Nfreq1-1]*(OctH-OctL)/(Nfreq1-1)); % Probe
                                                          %  freqs
f2of1=1.1;                      % suppressor:probe freq ratio f2/f1
f2mf1=round((1-f2of1)*freq1); % f2-f1;
nn1=round(1/(1-f2of1));       % round frequencies to nearest 1 Hz
nn2=nn1-1;
freq1=nn1*f2mf1;
```

```
freq2=nn2*f2mf1;
t0=clock;
for ifreq1=1:Nfreq1;
  freq=[freq1(ifreq1),freq2(ifreq1)]; % freqs of the two tones
  Nfreq=length(freq);                 % =2 (two tone suppression)
  str1= num2str(100*ifreq1/Nfreq1);
  str2= num2str(etime(clock,t0)/60); str2= str2(1:min(5,length(str2)));
  str3= num2str(freq(1));
  disp(' ');disp(' ');disp(' ');
  disp('IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII');
  disp('IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII');
  disp(...
      ['Freq= ',str3,' Hz; ',str1,' % complete after ',str2,' mins'])

  %
  %                    Derived & Preset Quantities
  %                    IIIIIIIIIIIIIIIIIIIIIIIII
  %
  pRef=20.e-6;
  Nfreq=length(freq);              % Number of frequencies
  WH=W*H;                          % csa of cochlea
  xs=L/(Nx-1);                     % Length step [m]
  x=[0:Nx-1].'*xs;                 % All NX points along the cochlea
  x2=x(1:Nx-1);                    % NX-1 points along CP (excl h/trema)
  len=1/kom;                       % place-freq length [m]
  omegaC0=sqrt(s0/m0);             % Char frequency at stapes [rad/s]
  omegaC=omegaC0*exp(-kom*x2);     % Char freq along BM [rad/s]
  r0=delta*sqrt(m0*s0);            % Damping/unit area at stapes [Ns/m^3]
  m1=m0;                           % Mass/unit area along CP [kg/m^2]
  s1=m1*omegaC.^2;                 % Stiffness/unit area along CP [N/m^3]
  r1=delta*sqrt(m1*s1);            % Damping/unit area along CP [Ns/m^3]
  YSrc=YSrc0*ones(size(freq));     % Source acoustic imp. at each freq
  ZEDnom=3.30e07;                  % Nominal eardrum acoustic impedance
  Pnom=10.^((SPLnom)/20.)*pRef;    % Nominal ear drum pressure
  QSrc=Pnom./ZEDnom;               % Source short circuit volume velocity
  ZHel=1e-12;  % Effective impedance at helicotrema (make v.small)
  %
  % === Two Port Representation of Ear Canal and Middle Ear ===
  %
  GOME=AEC/(ASt*GLME);  % Overall middle ear static pressure gain
  if MidEarFlg==0  % Transparent
    TEDoSt_11=ones(size(freq)) *1/GOME;  % <=== Stapes to Ear Drum
    TEDoSt_12=zeros(size(freq));
    TEDoSt_21=zeros(size(freq));
    TEDoSt_22=ones(size(freq)) *GOME;
  elseif MidEarFlg==3  % Kringlebotn
    % NB: in CGS units
    milli=1.e-3;  micro=1.e-6;
    La=1.00*milli;  Ca=3.90*micro;   Ra=60;
                    Ct=0.40*micro;
    Ld=7.50*milli;
    Ls=66.0*milli;  Cs=0.30*micro;   Rs=20;
                    Cr=1.30*micro;   Rr=120;
                    Cm=0.38*micro;   Rm=120;
    Lo=22.0*milli;                   Ro=20;
                    Ci=0.30*micro;   Ri=6000;
    Lc=46.0*milli;  Cc=0.56*micro;   Rc=330;

MidEarParam=[ASt,AEC,GLME,La,Ca,Ra,Ct,Ld,Ls,Cs,Rs,Cr,Rr,Cm,Rm,Lo,Ro,Ci,Ri
,Lc,Cc,Rc];
    [TEDoSt_11,TEDoSt_12,TEDoSt_21,TEDoSt_22]=t_kring(MidEarParam,freq);
  end;
  % === Ear Canal ===
  rhoc0=rho0*c0;
  kL=2*pi*freq*LEC/c0;
  sinkL=sin(kL);
  coskL=cos(kL);
  tankL=tan(kL);
  % === Transmission Matrices ===
  TECoED_11=coskL;                  % <=== Ear Drum to Ear Canal
```

```
TECoED_12=i*sinkL*rhoc0/AEC;
TECoED_21=i*sinkL*AEC/rhoc0;
TECoED_22=coskL;
TECoSt_11=TECoED_11.*TEDoSt_11+TECoED_12.*TEDoSt_21; % <=== Stapes
TECoSt_12=TECoED_11.*TEDoSt_12+TECoED_12.*TEDoSt_22; %        to
TECoSt_21=TECoED_21.*TEDoSt_11+TECoED_22.*TEDoSt_21; %     Ear Canal
TECoSt_22=TECoED_21.*TEDoSt_12+TECoED_22.*TEDoSt_22;
TSrcSt_11=TECoED_11.*TEDoSt_11+TECoED_12.*TEDoSt_21; % <=== Stapes
TSrcSt_12=TECoED_11.*TEDoSt_12+TECoED_12.*TEDoSt_22; %         to
TSrcSt_21=TECoED_21.*TEDoSt_11+TECoED_22.*TEDoSt_21; %      Source
TSrcSt_22=TECoED_21.*TEDoSt_12+TECoED_22.*TEDoSt_22;
% Refer Source Q and Y to the stapes.
QStSrc=QSrc./(TSrcSt_12.*YSrc+TSrcSt_22);                          % Q
YStSrc=(TSrcSt_11.*YSrc+TSrcSt_21)./(TSrcSt_12.*YSrc+TSrcSt_22); % Y

%              Cochlear Response Calculation
%              IIIIIIIIIIIIIIIIIIIIIIIIIIII

randn('seed',0);  %reset seed
if ScatFlg==1|ScatFlg==5     % Z & S broad band random scattering
  for iavg=1:irealiz
    inhomo=randn(size(x));
  end;
  inhomo=inhomo(1:Nx-1);
  Wn=2*xs/ScatLen;  % LP cut-off spatial freq /(1/2 sampling rate)
  if Wn<1.0
    [BB,AA]=butter(1,Wn);  % 1st order butterworth low pass filter
    inhomo=filtfilt(BB,AA,inhomo);
  end
elseif ScatFlg==2 % Strube periodic corrugations
  Kappa=1/ScatLen;
  inhomo=cos(2*pi*Kappa*x2);
elseif ScatFlg==3 % Strube narrow-band corrugations
  for iavg=1:irealiz
    inhomo=randn(size(x));
  end;
  inhomo=inhomo(1:Nx-1);
  Kappa=1/ScatLen;
  KappaBW=1/ScatLenBW;
  Wn=[Kappa-KappaBW/2,Kappa+KappaBW/2]*2*xs;
  if Wn<1.0
    [BB,AA]=butter(1,Wn);  %1st order butterworth bandpass filter
    inhomo=filtfilt(BB,AA,inhomo);
  end
elseif ScatFlg==4 % Point reflection site
  inhomo=0.5*(1+sign(x2-ScatLen)).*hamming(length(x2)).';
else % smooth CP
  inhomo=zeros(size(x2));
end
%
%              Single Tone Cochlear Response Calculation
%              IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
p=zeros(Nx,Nfreq);
vCP_tmp=zeros(Nx,Nfreq);
ZCP_QL_tmp=zeros(Nx,Nfreq);
ZOHC_QL_tmp=zeros(Nx,Nfreq);
u=zeros(Nx,Nfreq);
p_ifr=zeros(Nx,1);
rhs=zeros(Nx,1);
rnew=zeros(Nx,1);
a=zeros(Nx,1);
Ns=24;            % time samples samples per period of stimulus tone
itr_max=30;       % max iterations before time-out
tolabs=0.05;      % % convergence tolerance for magnitude
tolang=0.05*pi;   % % convergence tolerance for phase
str1= num2str(tolabs*100);
str2= num2str(tolang*180/pi);
%ncon=  zeros(itr_max,Nfreq+1);
%avgabs=zeros(itr_max,Nfreq+1);
%mxabs= zeros(itr_max,Nfreq+1);
```

```
%imxabs=zeros(itr_max,Nfreq+1);
%mxang= zeros(itr_max,Nfreq+1);
%imxang=zeros(itr_max,Nfreq+1);
disp('=O=O=O=O SINGLE ISOLATED TONES O=O=O=O=')
disp(' ')
disp(' convergence tolerances: ')
disp([' magnitude=',str1,'%,    phase=',str2,' deg'])
disp(' ')
% loop over the two stimulus tone frequencies
for ifr=1:Nfreq;
  str0= int2str(ifr);
  omega=2*pi*freq(ifr);
  iomega=i*omega;
  Tp=1/freq(ifr);                    % period [s]
  time=Tp*(-Ns/2:Ns/2-1)/Ns;         % time variable
  Mexp=diag(exp(-iomega*time))*2/Ns; % matrix to extract Fourier coeff.
  ZPass=s1/iomega+r1+iomega*m1;      % passive
  if ScatFlg~=5    % Scattering impedance based on damping
    ZScat=ScatSize*inhomo.*r1;
  else             % Scattering impedance based on stiffness
    ZScat=ScatSize*inhomo.*s1/iomega;
  end
  ZCP1=ZPass+ZScat;
  beta=omega./omegaC;
  Gxom=d0*(1+i*beta)./(deltaSC+i*(beta-sigma^2./beta));
  Gxom=[Gxom;0];
  ZOHC_MET=e0*[omegaC;0].*Gxom;  %the Z associated with the mech=>elec
%                   transduction (ie. linear part, called ZOHC by K & dB)
  phiZ=angle(ZOHC_MET);
  ZOHC_QL=ZOHC_MET;  % initialise quasilinear imp. to linear impedance
  ZCP1=[ZCP1;ZHel];
  ZCP_QL=ZCP1-ZOHC_QL;
  QStSrc_ifr=QStSrc(ifr);
  YStSrc_ifr=YStSrc(ifr);
  itr=0;
  loop=1;                          % setloop = true
  vCP=zeros(Nx,1);
  vCP_mid=vCP;
  vCP_old=vCP_mid;
  disp(...
[' freq : itn : max abs err @ worst elmnt : max ang err @ worst elmnt'])
  while (loop) % start iterative loop
    itr=itr+1;
    vCP_old=vCP_mid;
    vCP_mid=vCP;
    b=2*iomega*rho*xs^2./(H*ZCP_QL);
    A11=-1-b(1)-2*iomega*rho*xs*YStSrc_ifr/WH;
    A12=1;
    rhs(1)=-QStSrc_ifr*iomega*rho*xs/WH;
    rnew(1)=rhs(1)/A11;
    a(1)=-A12/A11;
    for ix=2:Nx-1;
      a(ix)=-1/(-2-b(ix)+a(ix-1));
      rnew(ix)=(rnew(ix-1)-rhs(ix))*a(ix);
    end;
    a(Nx)=-1/(-1-b(Nx)+a(Nx-1));
    rnew(Nx)=a(Nx)*(rnew(Nx-1)-rhs(Nx));
    p_ifr(Nx)=rnew(Nx);
    for ix=Nx-1:-1:1;
      p_ifr(ix)=rnew(ix)+a(ix)*p_ifr(ix+1); % fluid pressure
    end;
    vCP_new=(-2*p_ifr)./ZCP_QL;              % CP velocity
    vCP=(vCP_new+0.6065*vCP_mid+0.3679*vCP_old)/1.9744;  % 1st order
                                  %  lag relaxtion (seems to work)
    phiV=angle(vCP);
    % one-sided Fourier series expansion
    % x runs down a column; time runs along a row.
    phiIxt=(phiV+phiZ)*ones(size(time))+ones(size(x))*omega*time;
    A_pMET=abs(vCP.*Gxom)*ones(size(time));
    Ixt=A_pMET.*cos(phiIxt);
```

```
      p_OHC_NLt=e0*([omegaC;0]*ones(size(time))).*tanh(Ixt);
      % one-sided Fourier series coeff.
      % sum over time to extract 1st order component
      % NB time is along rows whilst "sum" sums over columns, so
      % therefore transpose
      p_OHC=sum(Mexp*p_OHC_NLt.').';
      ZOHC_QL=(p_OHC./vCP);
      ZCP_QL=ZCP1-ZOHC_QL;
      delta_abs=abs((vCP-vCP_mid)./vCP);
      delta_ang=abs(angle(vCP_mid./vCP));
      converged=all(delta_abs<tolabs) & all(delta_ang<tolang) & (itr>=3);
      % some convergence indicators:
      ncon(itr,ifr)=sum(delta_abs<tolabs)/Nx;  % proportion of converged
                                               % elements
      avgabs(itr,ifr)=sqrt(sum(delta_abs.*delta_abs)/Nx);  % average of
                                             % deltas over all elements
      [mxabs(itr,ifr),imxabs(itr,ifr)]=max(delta_abs);  % max error, and
                                                        %  its location
      [mxang(itr,ifr),imxang(itr,ifr)]=max(delta_ang);  % max error, and
                                                        % its location
      str1= sprintf('%4i',ifr);
      str2= sprintf('%3i',itr);
      str3= sprintf('%10.3f',mxabs(itr,ifr)*100);
      str4= sprintf('%11i',imxabs(itr,ifr));
      str5= sprintf('%6.0f',mxang(itr,ifr)*180/pi);
      str6= sprintf('%4i',imxang(itr,ifr));
disp ...
([' ',str1,' : ',str2,' : ',str3,'%   ',str4,' : ',str5,' deg ',str6])
      if converged
        loop=0;
        disp('  ')
        disp(['converged after ',num2str(itr),' iterations'])
      end
      if ~converged&itr>=itr_max
        disp('  ')
        disp(' *** WARNING *** ')
        disp(' failed to converge after max iterations ')
        loop=0
      end
    end;  % end of iterative loop
    itrmx(ifr)=itr;
    ZCP_QL_tmp(:,ifr)=ZCP_QL;
    ZOHC_QL_tmp(:,ifr)=ZOHC_QL;
    vCP_tmp(:,ifr)=vCP;                  % CP velocity
    u(2:Nx,ifr)=diff(p_ifr)/(-iomega*rho*xs);
    u(1,ifr)=u(2,ifr)-vCP(1)*xs/H;
    p(:,ifr)=p_ifr;
  end; % end loop over both stimulus tones
  clear vCP;vCP=vCP_tmp;clear vCP_tmp;
  clear ZCP_QL;ZCP_QL=ZCP_QL_tmp;clear ZCP_QL_tmp;
  clear ZOHC_QL;ZOHC_QL=ZOHC_QL_tmp;clear ZOHC_QL_tmp;
  % store isolated primaries
  u_pri=u;
  vCP_pri=vCP;
  p_pri=p;
  ZCP_QL_pri=ZCP_QL;
  ZOHC_QL_pri=ZOHC_QL;
  %
  %              Two Tone Cochlear Response Calculation
  %              IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
  %
  Ns0=24;              % min samples per stimulus tone period
  tolabs=0.01;         % convergence tolerance for magnitude
  tolang=2*pi/180;     % convergence tolerance for phase (in radians)
  str1= num2str(tolabs*100);
  str2= num2str(tolang*180/pi);
  disp(' ')
  disp('=O=O=O=O TWO TONE SUPPRESSION O=O=O=O=')
  disp(' ')
  disp(' convergence tolerances: ')
```

```
disp([' magnitude=',str1,'%,    phase=',str2,' deg'])
disp(' ')
itr_max=30;
omega=2*pi*freq;
Tp=1./freq;
Tp0=1/gcd(round(freq(1)),round(freq(2)));  % lowest common period [s]
Ns=Ns0*Tp0/Tp(2);                          % required number of samples
time=Tp0*(-Ns/2:Ns/2-1)/Ns;                % time variable
% initialise matrices: expand some vectors to two dimensions:
% =>ZCP1,ZOHC_MET,ZCP_QL,phiZ,vCP,vCP_new,vCP_mid,vCP_old
omega=2*pi*freq;  % [1 x 2]
iomega=i*omega;
ZPass=s1*(1./iomega)+r1*ones(size(omega))+m1*ones(size(x2))*iomega;
                                                        % size [Nx-1,2]
if ScatFlg~=5    % Scattering impedance independent of freq
  ZScat=ScatSize*(inhomo.*r1)*ones(size(omega));
else             % Scattering impedance dependent on freq
  ZScat=ScatSize*(inhomo.*s1)*(1./iomega);
end
ZCP1=ZPass+ZScat;
beta=(1./omegaC)*omega;  % size [Nx-1,2]
Gxom=d0*(1+i*beta)./(deltaSC+i*(beta-sigma^2./beta));
Gxom=[Gxom;[0,0]];
omegaC_2=[omegaC;0]; % size [Nx,1]
% the Z associated with the mech=>elec transduction (ie. linear part,
% called ZOHC by K & dB)
ZOHC_MET=e0*(omegaC_2*ones(size(omega))).*Gxom;  % size [Nx,2]
phiZ=angle(ZOHC_MET);        % size [Nx,2]
% initialise quasilinear impedance to value obtained above for
% isolated primaries
ZOHC_QL=ZOHC_QL_pri;
ZCP1=[ZCP1;[ZHel,ZHel]];                    % size [Nx,2]
ZCP_QL=ZCP1-ZOHC_QL;
vCP=vCP_pri;                  % initialise to isolated primary values
itr=0;                          % count iterations
loop=1;
vCP_mid=vCP_pri;
vCP_old=vCP_pri;
disp('  ')
disp(...
[' iter : freq : max abs err @ worst elmnt : max ang err @ worst elmnt'])
while (loop)
  itr=itr+1;
  vCP_old=vCP_mid;
  vCP_mid=vCP;
  b=(ones(size(x))*iomega).*(2*rho*xs^2./(H*ZCP_QL));  % size [Nx,2]
  for ifr=1:Nfreq % loop over two stimulus tone frequencies
    A11=-1-b(1,ifr)-2*iomega(ifr)*rho*xs*YStSrc(ifr)/WH;
    A12=1;
    rhs(1)=-QStSrc(ifr)*iomega(ifr)*rho*xs/WH;
    rnew(1)=rhs(1)/A11;
    a(1)=-A12/A11;
    for ix=2:Nx-1;
      a(ix)=-1/(-2-b(ix,ifr)+a(ix-1));
      rnew(ix)=(rnew(ix-1)-rhs(ix))*a(ix);
    end;
    a(Nx)=-1/(-1-b(Nx,ifr)+a(Nx-1));
    rnew(Nx)=a(Nx)*(rnew(Nx-1)-rhs(Nx));
    p_ifr(Nx)=rnew(Nx);
    for ix=Nx-1:-1:1;
      p_ifr(ix)=rnew(ix)+a(ix)*p_ifr(ix+1);   % fluid pressure
    end;
    p(:,ifr)=p_ifr;
  end
  vCP_new=(-2*p)./ZCP_QL;                     % CP velocity
  vCP=(vCP_new+0.6065*vCP_mid+0.3679*vCP_old)/1.9744;  % 1st order lag
                                                       %   relaxation
  phiV=angle(vCP);
  % one-sided Fourier series expansion
  % x runs down a column; time runs along a row.
```

```
    Ixt=zeros(Nx,Ns);
    for ifr=1:Nfreq % loop over two stimulus tone frequencies
      A_pMET=abs(vCP(:,ifr).*Gxom(:,ifr));
      phiIxt=(phiV(:,ifr)+phiZ(:,ifr))*ones(size(time))...
             +ones(size(x))*omega(ifr)*time;
      Ixt_ifr=(A_pMET*ones(size(time))).*cos(phiIxt);
      Ixt=Ixt+Ixt_ifr;
    end
    p_OHC_NLt=e0*(omegaC_2*ones(size(time))).*tanh(Ixt);
    % one-sided Fourier series coeff.
    % sum over time to extract 1st order component
    % NB time is along rows whilst "sum" sums over columns, so therefore
    % transpose
    for ifr=1:Nfreq % loop over two stimulus tone frequencies
      % Method 1: create [Ns,Ns] diagonal matrix M
      %           then matrix mult: [Ns,Ns]*[Ns,Nx]=[Ns,Nx]
      %           then sum & transpose
      % ==> [Ns,Ns] matrix for extracting Fourier coeff.
      % Mexp=diag(exp(-iomega(ifr)*time))*2/Ns;
      % p_OHC(:,ifr)=sum(Mexp*p_OHC_NLt.').';
      % Method 2: create [Ns,Nx] mx M (repeated cols)
      %           then array mult: [Ns,Nx].*[Ns,Nx]=[Ns,Nx} as before
      % ==> [Ns,Nx] matrix for extracting Fourier coeff.
      % NB method 2: much faster
      Mexp=(exp(-iomega(ifr)*time.')*2/Ns)*ones(size(x.'));
      p_OHC(:,ifr)=sum(Mexp.*p_OHC_NLt.').';
    end;
    ZOHC_QL=(p_OHC./vCP);
    ZCP_QL=ZCP1-ZOHC_QL;
    for ifr=1:Nfreq % loop over two stimulus tone frequencies
      delta_abs=abs((vCP(:,ifr)-vCP_mid(:,ifr))./vCP(:,ifr));
      delta_ang=abs(angle(vCP_mid(:,ifr)./vCP(:,ifr)));
      % some convergence indicators:
      conv(ifr)=all(delta_abs<tolabs) & all(delta_ang<tolang);
      % some convergence indicators:
      ncon(itr,2+ifr)=sum(delta_abs<tolabs)/Nx;  % proportion of
                                                 % converged elements
      avgabs(itr,2+ifr)=sqrt(sum(delta_abs.*delta_abs)/Nx); % average of
                                                % deltas over all elements
      [mxabs(itr,2+ifr),imxabs(itr,2+ifr)]=max(delta_abs);  % max overall
                                                   % error and its freq
      [mxang(itr,2+ifr),imxang(itr,2+ifr)]=max(delta_ang);  % max error,
                                                % and its location
      str1= sprintf('%3i',itr);
      str2= sprintf('%3i',ifr); ;
      str3= sprintf('%9.3f',mxabs(itr,2+ifr)*100);
      str4= sprintf('%11i',imxabs(itr,2+ifr));
      str5= sprintf('%6.0f',mxang(itr,2+ifr)*180/pi);
      str6= sprintf('%4i',imxang(itr,2+ifr));
      disp([' ',str1,'  : ',str2,'  : ',str3,'%   ',str4,...
      ' : ',str5,' deg    ',str6])
    end
    converged=conv(1)& conv(2)&(itr>=3);
    if converged
      loop=0;
      disp(' ')
      disp(['converged after ',num2str(itr),' iterations'])
    end
    if ~converged&itr>=itr_max
      disp(' ')
      disp(' *** WARNING *** ')
      disp(' failed to converge after max iterations ')
      loop=0
    end
  end;  % end of iterative loop
  clear A_pMET phiIxt Ixt;
  u(2:Nx,:)=diff(p).*(ones(Nx-1,1)*(1./(-iomega*rho*xs)));
  u(1,:)=u(2,:)-vCP(1,:)*xs/H;
  %
  %             External Responses Calculation
```

```
  %              IIIIIIIIIIIIIIIIIIIIIIIIIIIII
  pSt=2*p(1,:);
  QSt=u(1,:)*WH;
  uSt=QSt/ASt;
  ZSt=pSt/QSt;
  pED=TEDoSt_11.*pSt+TEDoSt_12.*QSt;
  QED=TEDoSt_21.*pSt+TEDoSt_22.*QSt;
  pEC_ifq=TECoED_11.*pED+TECoED_12.*QED;
  QEC_ifq=TECoED_21.*pED+TECoED_22.*QED;
  pEC(ifreq1,:)=pEC_ifq;  % pEC(iprobe freq,itone)
  QEC(ifreq1,:)=QEC_ifq;
  uED=QED/AEC;
  ZED=pED./QED;
  clear vCP_new vCP_mid vCP_old p_OHC_NLt Mexp Ixt_ifr
  ID=IDstart+ifreq1-1;
  fname=[matroot,int2str(floor(ID/100)),int2str(floor(rem(ID,100)/10))...
        ,int2str(rem(rem(ID,100),10)),'.mat'];
  varlist=' freq QSrc QEC pEC QED pED QSt pSt uSt YSrc TECoSt_11...
            TECoSt_12 TECoSt_21 TECoSt_22';
  command=['save ',fname,varlist];
  eval(command);
end %loop over freqs
```

Appendix III Parametric Model of SFOAE Frequency Functions

In this appendix, the theory behind the 3-parameter and 4-parameter models of the SFOAE frequency spectrum is discussed and the results quoted in sections 6.5 and 6.6 are derived. A Matlab program listing for the 4-parameter model is also presented.

III-1 3-parameter Model

This analysis shows how the three parameters in the model (i.e., the two Butterworth filter parameters, together with the RMS value of the signal) can be determined from the first few terms of the autocorrelation function of the random process. In the following discussion, it is useful to consider SFOAE frequency function sampled at equal η-intervals (i.e., log frequency intervals) as being the digital signal of interest. It also is useful to consider the η variable as analogous to time. Following this analogy, the terms *time* and *frequency* will be used to refer to η and the ϕ respectively. This allows the use of the familiar signal processing terms such as 'bandpass' or 'stationarity'.

As a first approximation, the SFOAE frequency function is modelled as a bandpass random signal which is stationary with respect to η. This can be thought of as arising from gaussian stationary white noise passed through a bandpass filter. In the 3-parameter model, a 2^{nd} order Butterworth filter has been chosen to fit the data. This is an ARMA (autoregressive, moving average) filter with nine filter coefficients. These nine coefficients are not independent, but instead are given by two filter parameters: the filter centre frequency and bandwidth. To obtain the spectrum of the data, the 3-parameter model then estimates these two filter parameters, plus the RMS value of the white noise.

This procedure is illustrated in the next section for a simpler case: a first order, bandpass Butterworth filter, which has only five filter coefficients instead of nine. Also for simplicity in this example the real form of the impulse response function is used instead of the analytic impulse response function used in the 3-parameter model.

III-2 The Autocorrelation Function for White Noise through a 1^{st} Order, Butterworth Bandpass Filter

Consider a signal arising from gaussian stationary white noise passed through a 1st order Butterworth bandpass filter. This is an ARMA filter with five filter coefficients and with the following input-output relationship:

$$x(n) = a_0 w(n) + a_1 w(n-1) + a_2 w(n-2) - b_1 x(n-1) - b_2 x(n-2) \qquad \text{[III–1]}$$

where $x(n)$ is the filter output sequence, $w(n)$ is the (white-noise) input sequence, and the a_n and b_n s are the MA and AR filter coefficients respectively.

These five coefficients are given by the two 1st order Butterworth filter parameters in the following relationships:

$$\begin{aligned}
a_0 &= \omega_{BW} / \left(1 + \omega_{BW} + {\omega_c}^2\right) \\
a_1 &= 0 \\
a_2 &= -a_0 \\
b_1 &= 2(\omega_c - 1) / \left(1 + \omega_{BW} + {\omega_c}^2\right) \\
b_2 &= \left(1 - \omega_{BW} + {\omega_c}^2\right) / \left(1 + \omega_{BW} + {\omega_c}^2\right)
\end{aligned} \qquad \text{[III–2]}$$

$$\begin{aligned}
\omega_c &= \sqrt{\omega_U \omega_L} \\
\omega_{BW} &= \omega_U - \omega_L \\
\omega_{U,L} &= \tan(\omega_{dU,dL} / 2 f_s)
\end{aligned}$$

where ω_{dU} and ω_{dL} are the desired upper and lower cut-off frequencies, f_s is the sampling rate, ω_U and ω_L are the dimensionless prewarped cut-off frequencies of the Butterworth filter, and ω_c and ω_{BW} are the dimensionless centre frequency and bandwidth.

By forming the autocorrelation function for the signal, x, equation [III–1] can be converted into the following relationship

$$\mathsf{R}_{xx}(m) = E\left[\, x(n)^{*}\, x(n+m) \,\right]$$

$$\mathsf{R}_{xx}(m) = \sum_{k=0}^{2} a_k \mathsf{R}_{wx}(k-m) - \sum_{k=1}^{2} b_k \mathsf{R}_{xx}(m-k) \quad \text{for all } m$$

and [III–3]

$$\mathsf{R}_{wx}(n) = \sigma_w^2 h(n)$$

$$h(n) = \sum_{k=0}^{2} a_k \delta(n-k) - \sum_{k=1}^{2} b_k h(n-k)$$

$$\delta(n) = \left\{ \begin{matrix} 1 \text{ for } n = 0 \\ 0 \text{ otherwise} \end{matrix} \right\}$$

where R_{xx} is the autocorrelation function of x, R_{wx} is the cross correlation function between the white noise and signal x, h is the impulse response function of the filter, δ is the unit impulse signal, σ_w is the RMS value of the white noise, $E[...]$ denotes the expectation operator and * denotes complex conjugation.

Expanding the summation in [III-3] explicitly gives:

$$\mathsf{R}_{xx}(0) = -b_1 \mathsf{R}_{xx}(1) - b_2 \mathsf{R}_{xx}(2) + \sigma_w^2 a_0 h(0) + \sigma_w^2 a_1 h(1) + \sigma_w^2 a_2 h(2)$$

$$\mathsf{R}_{xx}(1) = -b_1 \mathsf{R}_{xx}(0) - b_2 \mathsf{R}_{xx}(1) + \sigma_w^2 a_1 h(0) + \sigma_w^2 a_2 h(1)$$

$$\mathsf{R}_{xx}(2) = -b_1 \mathsf{R}_{xx}(1) - b_2 \mathsf{R}_{xx}(0) + \sigma_w^2 a_2 h(0)$$

$$\mathsf{R}_{xx}(3) = -b_1 \mathsf{R}_{xx}(2) - b_2 \mathsf{R}_{xx}(1)$$

$$\mathsf{R}_{xx}(4) = -b_1 \mathsf{R}_{xx}(3) - b_2 \mathsf{R}_{xx}(2)$$

..... [III–4]

$$\mathsf{R}_{xx}(m) = -b_1 \mathsf{R}_{xx}(m-1) - b_2 \mathsf{R}_{xx}(m-2) \qquad \text{for } m > 2$$

and

$$h(0) = a_0$$

$$h(1) = a_1 - b_1 h(0) = a_1 - b_1 a_0$$

$$h(2) = a_2 - b_1 h(1) - b_2 h(0) = a_2 - b_1 a_1 + {b_1}^2 a_0 - b_2 a_0$$

The first three rows can be further expanded by replacing the h's with a's and b's by using the last three equations in [III-4]. Thus, given the first five values of the autocorrelation function, $\mathsf{R}_{xx}(m)$, $m = 0, 1,...4$ in [III–4] it is possible to calculate the five ARMA filter

coefficients (though the AR filter coefficients cannot be separated from the unknown RMS value of the white noise input, σ_w). Note, however, that the first three rows of [III–4] contain nonlinear terms comprising products of a's with h's, which expand to give products of a's with b's. Thus, unlike Burg's method (1978a, 1978b) discussed later, the five ARMA filter coefficients would have to be solved using nonlinear methods. Having solved for the ARMA filter coefficients, the Butterworth filter parameters, ω_C and ω_{BW}, could be calculated (again using nonlinear methods) from the equations in [III–2]. Note that ω_C and ω_{BW} are overdetermined, since there are five (possibly independent) ARMA filter coefficients and only two Butterworth filter parameters. If the autocorrelation function in [III–4] truly arises from white noise through a 1st order Butterworth filter, then this is not a problem, since the ARMA filter coefficients will be interdependent such that all five equations in [III-2] are satisfied exactly. If this is not the case, however, then a least squares method becomes appropriate for inverting [III–2].

It should be noted that the purpose of the above analysis is to illustrate that all the information required to determine the unknown filter parameters is contained in only the first few terms of the autocorrelation function. It is not intended to give the actual solution method adopted in this thesis. This is discussed in the next section.

III-3 Solution Method for the 3-parameter Model

Unlike the analysis in the previous section, the 3-parameter model assumes a 2nd order, rather than a 1st order Butterworth filter. This has nine, rather than five ARMA filter coefficients (though is still has only two free filter parameters: ω_C and ω_{BW}). Also, in order to achieve a one-sided filter, the analytic form of the filter is used, rather than the real form. In fact, the two stage solution method described above (i.e., first calculating the ARMA filter coefficients from the autocorrelation function in [III–4], and then the filter parameters from the ARMA filter coefficients in [III–2]) is unnecessary. Instead, the following procedure has been adopted.

First the autocorrelation function is estimated from the (finite) signal, $\boldsymbol{x}(n)$, of length, N, by the biased estimator:

$$\tilde{\mathsf{R}}_{xx}(m) = \frac{1}{N}\sum_{n=0}^{N-m-1} \boldsymbol{x}^*(n)\boldsymbol{x}(n+m) \quad \text{for } m = 0...N-1 \qquad \text{[III–5]}$$

and normalised using the signal variance:

$$\tilde{K}_{xx}(m) = \frac{\tilde{R}_{xx}(m)}{\tilde{R}_{xx}(0)} \qquad \text{[III–6]}$$

The three unknown parameters, ω_{BW}, ω_c and σ_w, are then estimated using the following iterative procedure.

1. Guess initial values of the three parameters, ω_{BW}, ω_c and σ_w.
2. Calculate the nine ARMA filter coefficients from the 2nd order Butterworth filter equation, which is an extension of the 1st order filter equation [III–2].
3. Calculate the impulse response function of the real 2nd order Butterworth filter from the 2nd order version of [III–1] with the nine ARMA filter coefficients and a delta function as the input signal:

$$\left.\begin{aligned} h_R(n) &= 0 \\ h_R(n) &= a_n - b_1 h_R(n-1) - b_2 h_R(n-2) - b_3 h_R(n-3) - b_4 h_R(n-4) \\ h_R(n) &= -b_1 h_R(n-1) - b_2 h_R(n-2) - b_3 h_R(n-3) - b_4 h_R(n-4) \end{aligned}\right\} \begin{cases} n < 0 \\ 0 \le n \le 5 \\ n > 5 \end{cases} \qquad \text{[III–7]}$$

4. Calculate the analytic form of the impulse response function of the filter:

$$\begin{aligned} \boldsymbol{h} &\equiv h_R + i h_I \\ h_I &\equiv \mathrm{Hilb}\{h_R\} \end{aligned} \qquad \text{[III–8]}$$

5. Calculate the fitted normalised autocorrelation function of $x(n)$ from the impulse response function in [III–7]:

$$\begin{aligned} K_{xx}(m) &= K_{hh}(m) \\ K_{hh}(m) &\equiv \frac{R_{hh}(m)}{R_{hh}(0)} \\ R_{hh}(m) &\equiv \sum_{n=0}^{N-m-1} \boldsymbol{h}^*(n)\boldsymbol{h}\ (n+m) \quad \text{for } m = 0....N-1 \end{aligned} \qquad \text{[III–9]}$$

6. Calculate the mean squared error between the fitted and the estimated normalised autocorrelation functions, across the first M_{Trunc} values:

$$\varepsilon_{MS}^2 \equiv \frac{1}{M_{Trunc}} \sum_{m=0}^{M_{Trunc}} \left(\tilde{\mathsf{K}}_{xx}(m) - \mathsf{K}_{xx}(m)\right)^2 \qquad \text{[III–10]}$$

The choice of M_{Trunc} is discussed in the section III-5.

7. Iterate steps 1-5 above on the three variables to minimise the error in [III–10].

This procedure is equivalent to solving the 2nd order equivalents of [III–4] and [III–2] numerically using the first M_{Trunc} values in the equivalent to [III–4].

III-4 Comparison of the 3-parameter Model with Maximum Entropy Spectral Analysis (Burg, 1978a)

The 3-parameter model differs from the maximum entropy spectral analysis method due to Burg (1978a) in two main respects. Firstly, Burg's method uses only AR filter coefficients (denoted b_n). This means that the equivalent equation to [III–4] contains only linear terms and can simply be solved by linear matrix operations. Secondly, Burg's method does not constrain the filter to be any particular form. Instead, the order of the filter (i.e., the number of AR filter coefficients) is increased parametrically until some stopping criterion is reached. It is for these reasons that Burg's method is not suited to the work in this thesis. Being a more general spectral estimation method, Burg's method does not make use of the knowledge of the shape of the filter that has been gained from the cochlear modelling work, and consequently it contains many more free parameters (i.e., AR filter coefficients) that need to be determined than does the 3-parameter model.

III-5 The Optimal Value of M_{Trunc}

If the autocorrelation function of the signal were known exactly and if it arose from white noise passed through a 2nd order Butterworth, then it has been found that only the first two values of the autocorrelation function are required to solve exactly for the two Butterworth

filter parameter, f_C and f_{BW} in [III-2]. Any additional values of the autocorrelation function would be redundant, adding no new information. The mean square error term in this case would be identically zero. However, in reality the situation differs from this ideal case in three ways. Firstly, the autocorrelation function is only an estimate based on a finite length of signal. Note also that estimates of the autocorrelation function in equation [III–5] become less reliable at higher lag values, as they are based on fewer and fewer points. Secondly, the signal is contaminated with noise. Thirdly, the signal does not conform perfectly to stationary gaussian white noise through a 2nd order Butterworth filter. As a consequence of these factors, the optimal value of M_{Trunc} is > 2.

A theoretical treatment of the these effects is extremely difficult. There is one argument for choosing $M_{Trunc} = 9$. This would then allow nine ARMA filter coefficients to be calculated. From these, the two Butterworth filter parameters could be found by a least squares method (cf., the 1st order case in [III-4]). This does not, however, overcome the three probelms listed above. Thus instead of attempting further theoretical analysis, the following numerical approach was adopted. Using many signal realizations from both ideal 2nd order Butterworth processes and from the cochlear models, the performance of the 3-parameter model was measured in the presence of additive noise, and for different values of $M_{Trunc.}$ Simulations were performed with realistic values of the filter parameters, lengths of signal, and signal-to-noise ratios. The result of the numerical analysis was that the performance of the estimator varied little, provided that M_{Trunc} lay between about 10% and 50% of the signal length. In practice, a value of 12.5% was chosen for the analysis of both the cochlear models and the experimental data. This corresponds to about 11 points for the 91 measured points in the SFOAE signal, which is close to the value $M_{Trunc} = 9$ discussed above. The reason for defining M_{Trunc} as a proportion of the signal length, rather than as a number of points is explained below.

In the experiments, the SFOAE signal is resampled from 91 equispaced points in the frequency domain, to give 256 equispaced points in the η-domain. This effectively resamples the autocorrelation function, and complicates the argument in the following way. On resampling, the lag value of the Mth point of the original autocorrelation function is greater than the lag values of the Mth point of the resampled autocorrelation function. However, the resampling operation does not add any new information, and therefore the first M points in the resampled autocorrelation function contain less information than the first M points of the

original autocorrelation function. From this it can be shown that the number of points in the resampled autocorrelation function required to define the ARMA filter coefficient must be increased. Thus if the first 9 points of $R_{xx}(m)$ are required for the original 91 point signal, then the first 26 points are required when $R_{xx}(m)$ is calculated from a resampled signal of length 256 points. (The fact that the original signal was equispaced in frequency, whilst the resampled signal is equispaced in η does not does not materially alter the argument.) Thus for a given random process estimated from a given signal, it appropriate to define M_{Trunc} as a fraction of the number of points in the signal.

III-6 4-parameter Model

In this section, equation [6.12] is derived. In the 4-parameter model, the measured signal, $\boldsymbol{y}(n)$, can be thought of as arising from a two stage process. Firstly white noise is passed through a 2nd order Butterworth filter to give the signal, $\boldsymbol{x}(n)$, as for the 3-parameter model. Then this signal undergoes a nonlinear transformation to give $\boldsymbol{y}(n)$.

$$\boldsymbol{x}(n) = \boldsymbol{h}(n) \otimes w(n)$$

$$\boldsymbol{y}(n) = \frac{\boldsymbol{g}_0 \boldsymbol{x}(n)}{1 - \boldsymbol{r}_0 \boldsymbol{x}(n)} \qquad \text{[III–11].}$$

The relationship between the autocorrelation function of $\boldsymbol{y}(n)$ and the filter coefficients in $\boldsymbol{h}(n)$ is now more complicated than in [III-4]. The analysis proceeds as follows.

For small values of the product $\boldsymbol{r}_0 \boldsymbol{x}(n)$ (i.e. $\boldsymbol{r}_0 \boldsymbol{x}(n) < 1$) [III–11] can be expanded using the binomial theorem, and truncated to a the first two terms (further terms can be included if desired).

$$\boldsymbol{y}(n) \approx \boldsymbol{g}_0 \boldsymbol{x}(n)[1 + \boldsymbol{r}_0 \boldsymbol{x}(n) + \boldsymbol{r}_0^2 \boldsymbol{x}^2(n) + \boldsymbol{r}_0^3 \boldsymbol{x}^3(n)...]$$

Truncating to the first two terms:

$$\boldsymbol{y}(n) \approx \boldsymbol{A}\boldsymbol{x}(n) + \boldsymbol{B}\boldsymbol{x}^2(n) \qquad \text{[III–12]}$$

$$\boldsymbol{A} \equiv \boldsymbol{g}_0$$

$$\boldsymbol{B} \equiv \boldsymbol{g}_0 \boldsymbol{r}_0$$

The autocorrelation function of $\boldsymbol{y}(n)$ is then given by:

$$\mathsf{R}_{yy}(m) = E\left[\, \boldsymbol{y}(n)^* \boldsymbol{y}(n+m) \,\right]$$
$$\equiv \mathsf{R}_1(m) + \mathsf{R}_2(m) + \mathsf{R}_3(m) + \mathsf{R}_4(m)$$

where

$$\mathsf{R}_1(m) \equiv \boldsymbol{A}^2 \; E\left[\, \boldsymbol{x}(n)^* \boldsymbol{x}(n+m) \,\right] \qquad \text{[III–13]}$$
$$\mathsf{R}_2(m) \equiv \boldsymbol{B}^2 \; E\left[\, \boldsymbol{x}^2(n)^* \boldsymbol{x}^2(n+m) \,\right]$$
$$\mathsf{R}_3(m) \equiv \boldsymbol{AB} \; E\left[\, \boldsymbol{x}(n)^* \boldsymbol{x}^2(n+m) \,\right]$$
$$\mathsf{R}_4(m) \equiv \boldsymbol{AB} \; E\left[\, \boldsymbol{x}^2(n)^* \boldsymbol{x} \;\; (n+m) \,\right]$$

Now $\mathsf{R}_1(m)$ in [III–13] is simply proportional to the terms appearing in the 3-parameter model in [III–3] and [III–9]:

$$\mathsf{R}_1(m) \equiv \boldsymbol{A}^2 \; \mathsf{R}_{xx}(m) = \boldsymbol{A}^2 \; \sigma_w^2 \mathsf{R}_{hh}(m) \qquad \text{[III–14]}$$

where σ^2_w is the variance of the white noise signal.

The term $\mathsf{R}_3(m)$ can be found by expanding the convolution integral for $\boldsymbol{x}(n)$ in terms of the filter impulse response, $\boldsymbol{h}(n)$, and a white noise signal, $w(n)$.

$$\mathsf{R}_3(m) \equiv \boldsymbol{AB} \; E\left[\, x(n)^* x^2(n+m) \,\right]$$
$$= \boldsymbol{AB} \; E\left[\sum_{j_1} \boldsymbol{h}(j_1)^* w(n-j_1)^* \sum_{j_2} \boldsymbol{h}(j_2) w(n+m-j_2) \sum_{j_3} \boldsymbol{h}(j_3) w(n+m-j_3) \,\right]$$
$$= \boldsymbol{AB} \sum_{j_1} \sum_{j_2} \sum_{j_3} \boldsymbol{h}(j_1)^* \boldsymbol{h}(j_2) \boldsymbol{h}(j_3) E\left[\, w(n-j_1)^* w(n+m-j_2) w(n+m-j_3) \,\right]$$

[III–15]

where all summations are taken from $-\infty$ to $+\infty$. To evaluate the functions $\mathsf{R}_3(m)$, recall that the signals inside the expectation operator in [III–15] are all assumed to be gaussian, with zero mean. It can be shown that, for any three jointly gaussian random variables, X, Y, and Z of zero mean, the following holds (Deutsch, 1965):

$$E[XYZ] = 0 \qquad \text{[III–16]}$$

It follows from [III–15] and [III–16] that:

$$\mathsf{R}_3(m) = \mathsf{R}_4(m) = 0 \qquad \text{[III–17]}$$

Following a similar expansion to [III–15], it can be shown that $\mathsf{R}_2(m)$ in [III–13] becomes:

$$\begin{aligned}\mathsf{R}_2(m) &\equiv \boldsymbol{B}^2\, E\left[\, \boldsymbol{x}^2(n)^*\, \boldsymbol{x}^2(n+m)\,\right] \\ &= \boldsymbol{B}^2 \sum_{j_1}\sum_{j_2}\sum_{j_3}\sum_{j_4} \boldsymbol{h}(j_1)^*\, \boldsymbol{h}(j_2)^*\, \boldsymbol{h}(j_3)\boldsymbol{h}(j_4)\times \\ &\qquad E\left[\, w(n-j_1)^*\, w(n-j_2)^*\, w(n+m-j_3)w(n+m-j_4)\,\right]\end{aligned} \qquad \text{[III–18]}$$

For any four jointly gaussian random variables, W, X, Y, and Z of non-zero mean, the expectation of their product is given by [III–19] (Deutsch, 1965).

$$\begin{aligned}E[WXYZ] = {} & E[WX]E[YZ] + E[WY]E[XZ] + E[WZ]E[XY] \\ & - 2E[W]E[X]E[Y]E[Z]\end{aligned} \qquad \text{[III–19]}$$

Equation [III–19] allows the expectation in [III–18] to be expanded as a sum of autocorrelation functions. Recall that $w(n)$ is gaussian stationary white noise, whose autocorrelation function is therefore a delta function at zero lag. Thus the expectation in [III–18] becomes:

$$\begin{aligned}&E\left[\, w(n-j_1)^*\, w(n-j_2)^*\, w(n+m-j_3)w(n+m-j_4)\,\right] = \\ &\qquad \sigma_w{}^4[\delta(j_1-j_2)\delta(j_3-j_4) + \\ &\qquad\quad \delta(m+j_1-j_3)\delta(m+j_2-j_4) + \\ &\qquad\quad \delta(m+j_1-j_4)\delta(m+j_2-j_3)]\end{aligned} \qquad \text{[III–20]}$$

$$\delta(j-k) = \begin{cases} 1 & \text{for } j = k \\ 0 & \text{for } j \neq k \end{cases}$$

where σ_W is the RMS value of the white noise process. Substituting [III–20] into [III–18] and using the sifting property of the delta function gives:

$$\begin{aligned}\mathsf{R}_2(m) = {} & \sigma_w^4 \boldsymbol{B}^2 \sum_{j_2}\sum_{j_4} \boldsymbol{h}(j_2)^*\, \boldsymbol{h}(j_2)^*\, \boldsymbol{h}(j_4)\boldsymbol{h}(j_4) + \\ & \sigma_w^4 \boldsymbol{B}^2 \sum_{j_3}\sum_{j_4} \boldsymbol{h}(j_3-m)^*\, \boldsymbol{h}(j_4-m)^*\, \boldsymbol{h}(j_3)\boldsymbol{h}(j_4) + \\ & \sigma_w^4 \boldsymbol{B}^2 \sum_{j_4}\sum_{j_3} \boldsymbol{h}(j_4-m)^*\, \boldsymbol{h}(j_3-m)^*\, \boldsymbol{h}(j_3)\boldsymbol{h}(j_4) \\ = {} & \sigma_w^4 \boldsymbol{B}^2 \sum_{j_2} \boldsymbol{h}^2(j_2)^* \sum_{j_4} \boldsymbol{h}^2(j_4) + \\ & 2\sigma_w^4 \boldsymbol{B}^2 \sum_{j_3} \boldsymbol{h}(j_3-m)^*\, \boldsymbol{h}(j_3) \sum_{j_4} \boldsymbol{h}(j_4-m)^*\, \boldsymbol{h}(j_4)\end{aligned} \qquad \text{[III–21]}$$

Because $\boldsymbol{h}(n)$ is analytic it follows that:

$$\sum_j \boldsymbol{h}^2(j)^* = \sum_j \boldsymbol{h}^2(j) = 0 \qquad \text{[III–22]}$$

The demonstration of [III-22] is as follows. The product of any two analytic functions is also analytic, as argued later. Therefore $\boldsymbol{h}^2(n)$ is analytic. The sum of terms in [III–22] represents the zero-frequency (or d.c. offset) of $\boldsymbol{h}^2(n)$. But since $\boldsymbol{h}^2(n)$ is analytic, its zero-frequency term is zero. Therefore equation [III–22] holds if $\boldsymbol{h}(n)$ is analytic.

To show that the product of two analytic functions is itself analytic, consider the Fourier transform of the product. This yields the convolution of two transforms in the frequency domain. Since both transforms are one-sided (by the definition of analyticity), the convolution of the two one-sided transforms is itself one sided. Inverse Fourier transforming this one-sided function gives the desired product. Therefore the product of the two analytic functions must itself also be analytic.

Thus, by noting the result in [III–22], and recalling the definition in [III–9], it follows that [III–21] simplifies to:

$$\begin{aligned} \mathsf{R}_2(m) &= 2\sigma_w^4 \boldsymbol{B}^2 \left[\sum_{j_3} \boldsymbol{h}(j_3 - m)^* \boldsymbol{h}(j_3) \right]^2 \\ &= 2\sigma_w^4 \boldsymbol{B}^2 \left[\sum_j \boldsymbol{h}(j)^* \boldsymbol{h}(j+m) \right]^2 \\ &= 2\sigma_w^4 \boldsymbol{B}^2 \, \mathsf{R}_{hh}{}^2(m) \\ &= 2\boldsymbol{B}^2 \, \mathsf{R}_{xx}{}^2(m) \end{aligned} \qquad \text{[III–23]}$$

Substituting [III–23], [III–17] and [III–14] into [III–13] gives:

$$\begin{aligned} \mathsf{R}_{yy}(m) &= \boldsymbol{A}^2 \, \mathsf{R}_{xx}(m) + 2\boldsymbol{B}^2 \, \mathsf{R}_{xx}{}^2(m) \\ &= \boldsymbol{g}_0^2 \mathsf{R}_{xx}(m) + 2\boldsymbol{g}_o^2 \, \boldsymbol{r}_o^2 \mathsf{R}_{xx}{}^2(m) \end{aligned} \qquad \text{[III–24]}$$

This gives the first two terms in equation [6.12]. The third term in [6.12] (and higher order terms if desired) can be obtained in a similar way by including higher terms in the binomial expansion in [III–12]. In fact, the influence of the third order terms has generally been found to be very small. Equation [III–24] can then be used to fit the 4-parameter model to the estimated autocorrelation function, as described in section 6.6.

Note that there is an inconsistency in the above approach. The binomial expansion in [III–12] is only valid when $\boldsymbol{r}_0 \boldsymbol{x}(n) < 1$. However, since $\boldsymbol{x}(n)$ is assumed to be gaussian, it can take any value from $-\infty$ to $+\infty$. Thus, in theory, the binomial expansion becomes inapplicable

for a small proportion of the theoretical signal. In practice, however, this is not a problem. The physical signal, $\boldsymbol{x}(n)$ is, of course, not truly unbounded, and thus not truly gaussian.

III-7 Matlab Listing of 4-parameter Model

```
function ...
[phi,Syy4,SyyMeas,etaOctLag,Ryy4,RyyMeas,phiC,phiBW,alpha,beta,eps]...
         = param4(y,etaOct_s);
%
% Matlab Version 4.2
%
% B.Lineton 19-8-1999
%
% 4-parameter model: returns estimates of the four parameters together
%  with error values and the raw and fitted spectra
%
% inputs:
%   y= SFOAE frequency function sampled at equal log freq intervals
%   etaOct_s= the sampling interval in octaves
%
% outputs:
%   phi= the independent phi variable
%   Syy4= fitted 4 parameter phi-spectrum at phi values
%   Shh4= fitted 4 parameter phi-spectrum of
%   SyyMeas= measured raw phi-spectrum
%   etaOctLag= the independent lag octave frequency variable
%   Ryy4= fitted 4-parameter autocorrelation function
%   RyyMeas= measured raw autocorrelation function
%   phiC,phiBW,alpha,beta = estimates of four parameters
%   eps= final error (poorness-of-fit)

Ny=length(y);
etaOctLag  =[0:Ny-1]*etaOct_s;              % octave lag variable
phi_s=1/etaOct_s/log(2);                    % sampling rate in phi
phi=[0:Ny/2-1]*phi_s/Ny;

% remove mean, and find Hilbert transform pair
ym=mean(y);
yd=y-ym;
ydr=real(yd);
ydr2=imag(hilbert(imag(yd)));

% estimate autocorrelation function of y from
% cross correlation of real & Hilbert transform of imag parts
RyyMeas=xcorr(ydr,ydr2,'biased');
RyyMeas=RyyMeas(Ny:2*Ny-1);

% Obtain initial estimates of phiC & phiBW from fft over selected range:

win1=hanning(Ny).';
Fyr=fft(conj(detrend(ydr.*win1,0)),Ny);    % window 1
Fyr2=fft(conj(detrend(ydr2.*win1,0)),Ny); % window 1
SyyMeas=abs(Fyr.*Fyr2)/Ny/mean(win1.^2);  % estimate of raw phi-spectrum

% parametric fit over selected range
Nord=2;              % Butterworth filter order
NRPoF=round(Ny/8);  % number of points used in err calculation
% starting guesses: suffix 0
[mx,iphi_max]=max(SyyMeas)
phiC0=phi(iphi_max);
phiBW0=phiC0/2;
% Run at three different values of alpha1 & take the one with the lowest
final error:

alpha0_1=0;
[phiC1,phiBW1,alpha1,beta1,eps0_1,eps1]= ...
param4_1(Nord,NRPoF,RyyMeas,phi_s,phiC0,phiBW0,alpha0_1);
```

```
alpha0_2=0.25;
[phiC2,phiBW2,alpha2,beta2,eps0_2,eps2]= ...
param4_1(Nord,NRPoF,RyyMeas,phi_s,phiC0,phiBW0,alpha0_2);
alpha0_3=0.5;
[phiC3,phiBW3,alpha3,beta3,eps0_3,eps3]=...
param4_1(Nord,NRPoF,RyyMeas,phi_s,phiC0,phiBW0,alpha0_3);

[epsMin,IepsMin]=min([eps1,eps2,eps3]);
if IepsMin==1
  phiC=phiC1;
  phiBW=phiBW1;
  alpha=alpha1;
  beta=beta1;
  eps=eps1;
  eps0=eps0_1;
  alpha0=alpha0_1;
elseif IepsMin==2
  phiC=phiC2;
  phiBW=phiBW2;
  alpha=alpha2;
  beta=beta2;
  eps=eps2;
  eps0=eps0_2;
  alpha0=alpha0_2;
elseif IepsMin==3
  phiC=phiC3;
  phiBW=phiBW3;
  alpha=alpha3;
  beta=beta3;
  eps=eps3;
  eps0=eps0_3;
  alpha0=alpha0_3;
end;

% calculate fitted autocorrelation and phi-spectrum
Rhh4=RhhCalc(phiC,phiBW,Ny,Nord,phi_s);
sigmayy=sqrt(2*RyyMeas(1));
beta_sq=sigmayy^2/(24*abs(alpha)^6+6*abs(alpha)^4+2*abs(alpha)^2+1);
beta=abs(sqrt(beta_sq));
Shh4=2*real(fft(conj([0.5*Rhh4(1),Rhh4(2:Ny)])));
Shh4=Shh4(1:Ny/2);
Ryy4=abs(beta)^2*(24*abs(alpha)^6*Rhh4.^4+6*abs(alpha)^4*Rhh4.^3+2*abs(alpha
)^2*Rhh4.^2+Rhh4);
Syy4=2*real(fft(conj([0.5*Ryy4(1),Ryy4(2:Ny)])));
Syy4=Syy4(1:Ny/2);

% Recalculate analytic forms of the raw measured functions:
RyyMeas=2*conj(hilbert([fliplr(RyyMeas),RyyMeas(2:Ny)]));
RyyMeas=RyyMeas(Ny:2*Ny-1);
SyyMeas=4*SyyMeas(1:Ny/2);

return
```

```
% IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
% ========================================================================
% IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII

function [fc,fBW,alpha,beta,Err0,Err]= ...
         param4_1(Nord,NRyy,RyyMeas,fs,fc0,fBW0,alpha0);
%
% 4-parameter model: returns estimates of the four parameters together with
% the initial and final error values for a given autocorrelation function.
%
% inputs:
%   Nord= filter order (=2)
%   NRyy= number of points in Ryy overwhich to calc. the error
%   RyyMeas= the measured raw autocorrelation function
%   fs= samlping rate
%   fc0,fBW0,alpha0 = initial estimates of three parameters
%
% outputs:
%   fc,fBW,alpha = final estimates of three parameters
%   Err0, Err = initial and final error

Novr=length(RyyMeas);
sizeRyyd1=size(RyyMeas);
if sizeRyyd1(1)~=1;  RyyMeas=RyyMeas.';end;
ts=1/fs;

% === Optimise Parameters: alpha,beta,flow,fupp ===
% uses Matlab function "fmins" to minimise output of function "Par4Err0"

fupp0=0.5*sqrt(fBW0^2+4*fc0^2)+0.5*fBW0;  % convert centre freq & BW to
flow0=0.5*sqrt(fBW0^2+4*fc0^2)-0.5*fBW0;  % upper & lower cut-off freqs

FMIN_OPTIONS(1)=1;            % display parameters
FMIN_OPTIONS(3)=(1.e-4)/100; % tolerance
FMIN_OPTIONS(14)=2000;        % max iterations
ParamVec0=[flow0,fupp0,alpha0];

Err0=Par4Err0(ParamVec0,RyyMeas,Nord,NRyy,fs);  % initial error
ParamVec=fmins('Par4Err0',ParamVec0,FMIN_OPTIONS,[],RyyMeas,Nord,NRyy,fs);
Err=Par4Err0(ParamVec,RyyMeas,Nord,NRyy,fs);  % final error

flow=ParamVec(1);
fupp=ParamVec(2);
alpha =abs(ParamVec(3));

sigmayy=sqrt(2*RyyMeas(1));
beta_sq=sigmayy^2/(24*abs(alpha)^6+6*abs(alpha)^4+2*abs(alpha)^2+1);
beta=abs(sqrt(beta_sq));
fBW=fupp-flow;
fc=sqrt(fupp*flow);

return
```

```
% IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
% ========================================================================
% IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII

function Err=Par4Err0(ParamVec,RyyMeas,Nord,NRhh,fs);
%
% calculates error between a non-linear bandpass npoise signal and
% model of noise signal based on:
% y= A1*x + A2*x^2
% where x is 1st order Butterworth band pass noise with a std of 1;
% ParamVec=[flow,fupp,alpha];
% flow, fupp=3dB Upper and Lower Cut-offs for B'worth filter, both in cycles
%            per unit time
% alpha=A2/A1;
% NRhh= points in Rhh to be included;
% fs= sampling rate in cycles per unit time.
% use fact that real & imag parts are related by the Hilbert transform

alpha_Limit=1;
Novr=length(RyyMeas);
sigmayy2=2*RyyMeas(1);
RyyMeas=2*conj(hilbert([fliplr(RyyMeas),RyyMeas(2:Novr)]));  % create full
Ryy from real part;
KyyMeas=RyyMeas(Novr+(0:NRhh-1))/sigmayy2; % normalised & truncated to NRhh
points

flow=ParamVec(1);
fupp=ParamVec(2);
alpha =abs(ParamVec(3));
alpha =min([alpha,alpha_Limit]);

small=fs/Novr;  % =fs/Novr; % ensure filter turning points are sensible
if flow <=small           flow=small;   end;
if fupp <=2*small         fupp=2*small; end;
if fupp >=fs/2-small      fupp=fs/2-small; end;
if flow >=fupp-small flow=fupp-small; end;

[B,A]=butter(Nord,[flow,fupp]/(fs/2));

symclick=[zeros(1,Novr),1,zeros(1,Novr-1)];  % click
htemp1=filter(B,A,symclick);                 % hr(t) (real IRF)
htemp2=filter(B,A,conj(fliplr(htemp1)));     %
Rhrhr=fliplr(htemp2);                        % Rhrhr= hr(t) (x) hr(-t)
Rhh=2*conj(hilbert(Rhrhr));                  % analytic Rhh (2-sided in
time)
Rhh=Rhh(Novr+1:2*Novr);                      % analytic Rhh (1-sided in
time)
Rhh=Rhh(1:NRhh);                             % truncate to NRhh points
Rhh=(Rhh)/real(Rhh(1));                      % Normalise
Kyy4=(24*alpha^6*Rhh.^4+6*alpha^4*Rhh.^3+2*alpha^2*Rhh.^2+Rhh)...
        /(24*alpha^6+6*alpha^4+2*alpha^2+1);  % 4-paramter fit to Ryy
Diff=KyyMeas-Kyy4;
Err=sum(abs(Diff).^2)/NRhh;

return
```

```
% IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
% ========================================================================
% IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII

function Khh=RhhCalc(fc,fBW,N,Nord,fs);
% Khh=RhhCalc(fc,fBW,N,Nord,fs);
% return the normalised autocorrelation function corresponding to
% a Butterworth filter
% inputs:
%   fc, fBW= centre freq & bandwidth in cycles per unit time
%   N= Number of points;
%   fs= sampling rate in cycles per unit time;

fupp=0.5*sqrt(fBW^2+4*fc^2)+0.5*fBW;
flow=0.5*sqrt(fBW^2+4*fc^2)-0.5*fBW;

[B,A]=butter(Nord,[flow,fupp]/(fs/2));

symclick=[zeros(1,N),1,zeros(1,N-1)];     % click
htemp1=filter(B,A,symclick);              % hr(t) (real IRF)
htemp2=filter(B,A,conj(fliplr(htemp1)));
Rhrhr=fliplr(htemp2);                     % Rhrhr= hr(t) (x) hr(-t)
Rhh=2*conj(hilbert(Rhrhr));               % analytic Rhh (2-sided in time)
Rhh=Rhh(N+1:2*N);                         % analytic Rhh (1-sided in time)
Khh=(Rhh)/real(Rhh(1));                   % normalised Rhh

return
```