University of Southampton
Faculty of Engineering, Science and Mathematics
Institute of Sound and Vibration Research

Modelling the Human Cochlea

Emery Mayon Ku

Fulbright Scholar

A thesis submitted for the degree of

Doctor of Philosophy

2008

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING, SCIENCE AND MATHEMATICS

INSTITUTE OF SOUND AND VIBRATION RESEARCH

Doctor of Philosophy

Modelling the Human Cochlea

by Emery Mayon Ku

One of the salient features of the human cochlea is the incredible dynamic range it possesses—the loudest bearable sound is 10,000,000 times greater than the softest detectable sound; this is in part due to an active process. More than twelve thousand hair-like cells known as outer hair cells are believed to expand and contract in time to amplify cochlear motions. However, the cochlea's response is more than just the sum of its parts: the local properties of outer hair cells can have unexpected consequences for the global behaviour of the system. One such consequence is the existence of otoacoustic emissions (OAEs), sounds that (sometimes spontaneously!) propagate out of the cochlea to be detected in the ear canal.

In this doctoral thesis, a classical, lumped-element model is used to study the cochlea and to simulate click-evoked and spontaneous OAEs. The original parameter values describing the microscopic structures of the cochlea are re-tuned to match several key features of the cochlear response in humans. The frequency domain model is also recast in a formulation known as state space; this permits the calculation of linear instabilities given random perturbations in the cochlea which are predicted to produce spontaneous OAEs. The averaged stability results of an ensemble of randomly perturbed models have been published in [(2008) 'Statistics of instabilities in a state space model of the human cochlea,' J. Acoust. Soc. Am. **124**(2), 1068-1079]. These findings support one of the prevailing theories of SOAE generation.

Nonlinear simulations of OAEs and the model's response to various stimuli are performed in the time domain. Features observed in the model include the saturation of the forces generated by the OHCs, compression of amplitude growth with increasing stimulus level, harmonic and intermodulation distortion, limit cycle oscillations that travel along the cochlear membranes, and the mutual suppression of nearby linear instabilities.

Table of Contents

LIST OF FIGURES ..... **V**

LIST OF TABLES ..... **XII**

LIST OF ACCOMPANYING MATERIAL ..... **XII**

LIST OF SYMBOLS ..... **XV**

LIST OF ABBREVIATIONS ..... **XIX**

1 INTRODUCTION **1**

1.1 THE HUMAN COCHLEA 1

1.1.1 Biology and measurements *2*

1.1.2 the Dominant source of cochlear nonlinearity *9*

1.2 OTOACOUSTIC EMISSIONS 9

1.2.1 SOAEs *11*

1.2.2 CEOAEs *12*

1.2.3 Otherwise-evoked OAEs *14*

1.3 MODELS OF THE COCHLEA 14

1.3.1 the Neely and Kim (1986) model *16*

1.3.2 Causes of variations in feedback gain *17*

1.4 AIMS AND THESIS STRUCTURE 17

1.5 CONTRIBUTIONS 18

2 THE NEELY AND KIM (1986) MODEL **20**

2.1 MODEL OVERVIEW 20

2.2 PASSIVE MICROMECHANICS 23

2.2.1 Variation of model parameters with longitudinal position *24*

2.2.2 Micromechanical frequency response functions *26*

2.2.3 Passive admittance *28*

2.2.4 Undamped natural frequencies *29*

2.3 ACTIVE MICROMECHANICS 30

2.3.1 Active admittance *31*

2.3.2 Micromechanical stability analysis *33*

2.4 COCHLEAR MACROMECHANICS 38

2.4.1 Numerical implementation of frequency domain macromechanics *39*

2.4.2 Response of the coupled cochlea *41*

3 STATE SPACE FORMULATION **51**

3.1 FORMULATION OVERVIEW 52

3.2 ISOLATED MICROMECHANICS 53

3.2.1 Stability: poles and zeros *56*

3.3 COUPLED MACROMECHANICS 64

3.3.1 Stability: uniform variation of $\gamma(x)$ *67*

3.3.2 Stability: step change in γ(x) *68*
3.3.3 Stability: random variation of γ(x) *69*
3.4 FREQUENCY DOMAIN RESPONSES 71
3.4.1 Input and output impedances *72*
3.5 TIME DOMAIN SIMULATIONS 75
3.5.1 Linear responses *75*
3.5.2 Nonlinear responses *78*
4 SPONTANEOUS OTOACOUSTIC EMISSIONS 87
4.1 THEORIES OF SOAE GENERATION 88
4.1.1 Theoretical predictions *91*
4.2 LINEAR STABILITY GIVEN PERTURBATIONS IN Γ(X) 92
4.2.1 Step changes in γ(x) *93*
4.2.2 Sinusoidal variations in γ(x) *97*
4.2.3 Random variations in γ(x) *102*
4.3 NONLINEAR TIME DOMAIN SIMULATIONS OF UNSTABLE COCHLEAE 108
4.3.1 Step changes in γ(x) *108*
4.3.2 Random variation in γ(x) *115*
4.4 DISCUSSION 124
4.4.1 Linear results *125*
4.4.2 Nonlinear results *126*
5 CLICK EVOKED OTOACOUSTIC EMISSIONS 137
5.1 LINEAR SIMULATIONS 140
5.1.1 BM Responses *140*
5.1.2 CEOAEs *145*
5.2 NONLINEAR SIMULATIONS 151
5.2.1 BM Responses *151*
5.2.2 CEOAEs *155*
5.2.3 Inhomogeneities in δ(x) *159*
5.3 DISCUSSION 161
6 CONCLUSIONS AND SUGGESTIONS FOR FURTHER WORK 166
6.1 STRENGTHS OF THE STATE SPACE MODEL 166
6.1.1 Cochlear responses *167*
6.1.2 SOAEs *168*
6.1.3 CEOAEs *170*
6.2 WEAKNESSES OF THE STATE SPACE MODEL 171
6.2.1 Cochlear responses *171*
6.2.2 OAEs *174*
6.3 SUGGESTIONS FOR FUTURE WORK 175
6.3.1 Further study and tuning of the model *175*
6.3.2 Calculations and simulations *177*

6.3.3 Wider topics of research *178*

A MIDDLE EAR AND EAR CANAL MODEL **180**

A.1 TWO-PORT NETWORKS 181

A.1.1 Calculating input impedances in a two-port network *182*

A.1.2 Independent responses of the ear canal and middle ear models *183*

A.2 RESPONSE OF THE COMBINED MIDDLE EAR AND EAR CANAL MODELS 191

A.2.1 Forward transfer function in frequency *192*

A.2.2 Forward transfer function in time *193*

A.2.3 Reverse transfer function in frequency *194*

A.2.4 Reverse transfer function in time *195*

A.3 VALIDATION OF THE TWO-PORT MODEL 196

A.4 REVERSE INPUT ADMITTANCE IN STATE SPACE 198

A.5 SUMMARY AND DISCUSSION 199

B DERIVATION OF NEELY AND KIM QUANTITIES **201**

B.1 BM AND TM ADMITTANCE 201

B.2 UNDAMPED NATURAL FREQUENCIES 203

B.3 FREQUENCY DOMAIN STABILITY CRITERION 205

B.4 BM AND TM EQUATIONS OF MOTION 207

C JOURNAL PUBLICATIONS **209**

C.1 ELLIOTT ET AL. 2007 209

C.2 KU ET AL. 2008 223

D GROWTH CURVES OF CEOAES **236**

E SUPPRESSION OF VAN DER POL OSCILLATOR SIMULATIONS **240**

E.1 THE VAN DER POL EQUATION 240

E.2 DISCUSSION 243

LIST OF REFERENCES **245**

List of Figures

Figure 1.1: The cochlea in relation to the outer and middle ears 2
Figure 1.2: Schematic representation of the uncoiled cochlea.....3
Figure 1.3: Schematized illustration of the BM travelling wave and fluid flow given sinusoidal excitation at the stapes 3
Figure 1.4: Cross-section of a single turn of the cochlea 4
Figure 1.5: Detailed view of the organ of Corti 4
Figure 1.6: Input-output functions of an IHC and an OHC from the base of a guinea pig cochlea 5
Figure 1.7: The upper surface of a chinchilla organ of Corti 6
Figure 1.8: Nonlinear growth and compression of the BM response 8
Figure 1.9: Variation in the amplitude of the BM response with level 8
Figure 1.10: Example of a SOAE measurement in a human 11
Figure 1.11: Example of a typical CEOAE measurement in a normal-hearing human subject obtained with commercially available equipment 13

Figure 2.1: Neely & Kim's (1986) micromechanical model of the cochlea 21
Figure 2.2: Schematic representation of Neely and Kim's 1-D macromechanical model of the cochlea 23
Figure 2.3.a-c: Variation of model parameters with position 26
Figure 2.4.a-d: Magnitudes and phases of the damped and near-undamped system admittance at x = 20.48 mm.....28
Figure 2.5.a-b: Illustration of the relative magnitudes and directions of motion of the BM and the TM at the first and second modes of oscillation. 29
Figure 2.6: Undamped, micromechanical BM and TM resonant modes as a function of position along the BM 30
Figure 2.7.a-d: Active BM admittance magnitude and phase as a function of position along the cochlea given a 1 kHz excitation 32

Figure 2.8.a-d: Active BM admittance magnitude and phase for the isolated micromechanical model as a function of frequency at x = 20.48 mm 33
Figure 2.9: General diagram of a feedback loop containing a micromechanical controller. The dotted lines represent the system input and output paths 34
Figure 2.10: Plot of the open-loop transfer function, *GH*, for the micromechanical oscillators at several positions along the cochlea 35
Figure 2.11: Isolated enhancement given $\gamma = 1$ and minimum values of gain, γ, that lead to instability for various positions along the cochlea 36
Figure 2.12: Calculated admittance of an isolated micromechanical element as a function of frequency at x = 20.48 mm for 3 values of γ 37
Figure 2.13: Schematic of a 1D box model of cochlear macromechanics: detail view of the conceptual construction of the macromechanics 38
Figure 2.14.a-d: Pressure and BM velocity magnitude and phase given a 1kHz stimulus tone in the coupled cochlea given active and passive models. 41
Figure 2.15.a-d: Pressure and BM velocity magnitude and phase as a function of frequency at the 1kHz characteristic place in the coupled cochlea given active and passive models. 42
Figure 2.16: Enhancement of TW motion in the coupled cochlea 43
Figure 2.17: Frequency-to-place map of the model compared to the Greenwood function for a human 44
Figure 2.18.a-b: Wavelength of the TW as a function of a) position for a 1 kHz tone, and b) frequency at the 1 kHz characteristic place 45
Figure 2.19: Wavelength of the TW at its peak as a function of position along the cochlea 45
Figure 2.20: Real and imaginary components of the wavenumber for the present model ..47
Figure 2.21: Distance to the characteristic place as a function of cochlear delay 48
Figure 2.22: Total TW phase accumulation from the base to the characteristic place as a function of frequency 48
Figure 2.23: Variation of cochlear response given several values of *N* 49

Figure 3.1: Illustration of the effect of poles upon the unforced system response 57
Figure 3.2.a-c: Poles and zeros of Neely and Kim's micromechanical model with $\gamma = 1$. ..59
Figure 3.3: Active and passive admittance of the BM at x = 20.48 mm. Also superimposed are the pole and zero in this frequency region 60

Figure 3.4.a-b: Plot of the zeros and poles of the micromechanical admittance at 20.48 mm along the cochlea given variations in feedback gain, γ 60
Figure 3.5.a-b: Plot of isolated, active micromechanical poles at four positions along the cochlea 63
Figure 3.6.a-b: Poles for all the isolated micromechanical elements plot together at one gain given passive and active models. 63
Figure 3.7.a-d: Stability plots of the coupled state space model.....68
Figure 3.8.a-d: Stability of two coupled cochlear models given step changes in γ 69
Figure 3.9.a-d: Stability of two models given smooth and rough distributions of feedback gain 70
Figure 3.10: Response of the model calculated using the state space formulation at 4 kHz compared with the frequency domain formulation 72
Figure 3.11.a-d: Comparison of Z'_{in} and Z_{out} magnitude and phase.....74
Figure 3.12: The first 20 ms of stapes acceleration that serves as an input to the time domain simulation 76
Figure 3.13: Magnitude and phase of a 3 kHz stimulus tone plotted as a function of position along the cochlea 77
Figure 3.14: The BM velocity in a linear cochlea given a windowed 3 kHz stimulus tone at 0 dB SPL 78
Figure 3.15.a-b: Boltzmann function characteristics: output vs. input, slope of output vs. input 79
Figure 3.16: Nonlinear saturation point as a function of position 80
Figure 3.17.a-b): Mesh of the first 30 ms of BM velocity given a 3 kHz input tone at 45 dB SPL and 90 dB SPL in a nonlinear cochlea. 81
Figure 3.18: Growth of the steady state BM response given increasing amplitudes at 3kHz in 3 dB steps 82
Figure 3.19.a-d: Normalised growth curve of the 1.5 kHz component, the 3 kHz component, the 6 kHz component, and the 9 kHz component of the BM response as a function of position along the cochlea 83
Figure 3.20.a-c: Growth of distortion components generated at the 3 kHz place; near (±~1 mm about) the best places of the distortion frequencies; and at the 6 kHz place.....84

Figure 4.1.a-b: The length scale and wavelength of the TW at its peak, and the predicted $f/\Delta f$ 92

Figure 4.2: Stability plot for a cochlea with its stepped gain as a function of position inset 93
Figure 4.3.a-d: Magnitude and phase of BM velocity for excitation at several frequencies given a baseline model with nominal gain, $\gamma(x) = 1$ 95
Figure 4.4: Stability plot for several cochlear models with varying values of *N*, the number of micromechanical elements. A stepped gain as a function of position was applied to each model and is inset in the figure 96
Figure 4.5.a-h: Stability plots models with 8 different values of λ_{sin} 98
Figure 4.6.a-b: Average undamping ratio of the instabilities and total number of unstable frequencies given variations in the sinusoidal wavelength of $\gamma(x)$. 99
Figure 4.7.a-h: Average undamping ratio of the instabilities that fall within half-octave bands as λ_{sin} is varied 100
Figure 4.8.a-h: Total count of the instabilities that fall within half-octave bands as λ_{sin} is varied 101
Figure 4.9.I-IV.a-d: The collected results from 4×200 cochlear models with randomly generated gain distributions 103
Figure 4.10.a-d: Plot of spacings between adjacent instabilities given various $\lambda_{cut\text{-}off}$ as a function of frequency for the results collected in Figure 4.9.....106
Figure 4.11.a-b: Average undamping ratio and number of unstable frequencies per cochlear model given variations in $\lambda_{cut\text{-}off}$ and the peak-to-peak variations in $\gamma(x)$ 107
Figure 4.12.a-b: Stability plot of a model with a stepped gain distribution 109
Figure 4.13.a-b: Mesh of results from a nonlinear unstable cochlea, simulated in the time domain 109
Figure 4.14.a-c: The first 60 ms of BM velocity at three locations in an unstable cochlear model 111
Figure 4.15.a-d: Superimposed stability plots of linear system and the spectrum of the pressure at the base of the nonlinear cochlea given four time windows 112
Figure 4.16.a-c: Variation of the magnitudes of various frequency components of the pressure at the base of an unstable cochlea with time 113
Figure 4.17.a-b: The magnitude and phase of the BM velocity at the linearly unstable frequency and its first two harmonics calculated as a function of position.....114
Figure 4.18: A windowed-perturbed distribution of $\gamma(x)$.....116
Figure 4.19: Stability of a cochlear model given the windowed-perturbed gain distribution as presented in Figure 4.18. 116

Figure 4.20: Mesh of BM velocity in the first 60 ms of a nonlinear simulation of an unstable cochlear model 117
Figure 4.21: The first 60 ms of BM velocity at three different locations following a click stimulus at the base.....118
Figure 4.22: Simultaneous plot of linear system stability and the pressure spectrum at the base of the cochlea at steady state 119
Figure 4.23.a-d: Simultaneous plot of linear system stability and the spectrum of the pressure at the base of the cochlea given four time windows 121
Figure 4.24.a-c: Variation of the magnitudes of various frequency components of the pressure at the base with time in an unstable cochlea 122
Figure 4.25.a-b: The pressure spectrum in panel (a) displays a number of selected limit cycles; these frequencies are used to compute distances between the limit cycle oscillations, as shown in panel (b) 124
Figure 4.26: Steady state wave velocity of three limit cycle frequency components along the nonlinear BM 128
Figure 4.27.a-b: Steady state magnitude and phase of the BM response as a function of position at the unstable frequency 129
Figure 4.28: Real part of the BM admittance as a function of position along the cochlea, given f = 1.214 kHz 130
Figure 4.29.a-b: Phase of 1.214 kHz TW and its reflected response at the location of the discontinuity 132
Figure 4.30.a-b: Steady state magnitude and phase of the $2f_0$ harmonic response of the unstable frequency 133
Figure 4.31: Real part of the admittance of three linearly unstable frequencies.....134

Figure 5.1.a-d: Response of linear cochlear models due to a 100 μs pulse of volume displacement equivalent to 20 dB SPL 142
Figure 5.2: Stability plot for a cochlear model with dense, random 0.75% peak-to-peak inhomogeneities in $\gamma(x)$.....143
Figure 5.3.a-d: Normalised BM displacement at 1.33 mm and 13.96 mm in a baseline linear cochlea due to a standard click with variations in gain 144
Figure 5.4.I-II.a-d: The response of the baseline active cochlea given a standard click at 20 dB SPL compared with the perturbed, stable model of Figure 5.2 146
Figure 5.5: The Morlet mother wavelet 148

Figure 5.6.a-d: Directly determined CEOAE waveform at the stapes and in the ear canal 149
Figure 5.7.a-b: Predicted CEOAE latencies for an active baseline cochlea compared to calculated, band-averaged latencies from simulated DD CEOAEs in perturbed cochleae 150
Figure 5.8.a-d: Nonlinear cochlear response of model presented in Figure 5.2 due to a standard click at 4 stimulus levels 152
Figure 5.9.a-f: Displacement at 1.33 mm and 13.96 mm in a baseline cochlea due to a standard click with variations in stimulus level from 0 dB to 120 dB SPL; the locations of the peaks given the same progression from linear to saturating; and the magnitude of the displacement at the peak 153
Figure 5.10.a-d: Nonlinear growth of the baseline active model: BM displacement at 4 positions along the BM 154
Figure 5.11.a-b: Directly-determined, normalised CEOAE pressure response from the nonlinear model at the stapes and in the ear canal 156
Figure 5.12.a-b: Detailed view of the normalised directly-determined CEOAE pressure response at the stapes and in the ear canal 156
Figure 5.13.a-d: CWT-derived latencies in 11 frequency bands at all stimulus levels 158
Figure 5.14.a-d: Normalised derived nonlinear CEOAE response at the stapes and in the ear canal 159
Figure 5.15.a-b: Perturbed saturation point as a function of position 160
Figure 5.16.a-b: Directly-determined CEOAE in the nonlinear model at the stapes and ear canal due to perturbations in $\delta(x)$ 161

Figure A.1 The human auditory system 180
Figure A.2: Two-port network representations of the ear canal and middle ear 181
Figure A.3: Two-port network representation given a load termination at the output port 182
Figure A.4: Schematic illustration of the ear canal model. 183
Figure A.5: Two-port network representation of the ear canal. 184
Figure A.6.a-d: Ear canal admittance as seen from the eardrum given blocked, lossy, flanged and unblocked terminations. 186
Figure A.7: Human middle ear 187
Figure A.8: Two-port network representation of the middle ear. 187
Figure A.9: Diagram of the middle ear impedances in the model of Kringlebotn (1988). 188

Figure A.10 Simplified block diagram of the network model of the middle ear. 190
Figure A.11.a-b: Input admittance of the middle ear as seen from the stapes given blocked and unblocked conditions at the eardrum. 191
Figure A.12: Two-port network representation of the combined middle ear and ear canal. 191
Figure A.13: Magnitude and phase of the forward transfer function frequency response: stapes acceleration given a volume displacement in the ear canal 193
Figure A.14: Standard click input to the model: stapes acceleration due to an ear canal volume displacement 194
Figure A.15: Reverse pressure frequency response: pressure at the outer ear per unit input pressure at the stapes 195
Figure A.16: Reverse pressure transfer function: pressure at the outer ear per unit input pressure at the stapes. 196
Figure A.17.a-b: Comparison of forward and reverse pressure gain functions 197
Figure A.18.a-b: Comparison of round-trip middle ear gain as produced by the two-port model and as experimentally measured by Puria (2003) 198
Figure A.19.a-b Comparison of model and measured reverse middle ear impedance frequency response 199

Figure B.1: General diagram of micromechanical controller. 205
Figure B.2.a-b: Force diagram for the BM and the TM in Neely and Kim's (1986) micromechanical model 207

Figure D.1.a-d: Growth curves of CEOAE amplitude in the first ms, and the last 20 ms 237
Figure D.2.a-b: Growth curves of CEOAE amplitude in short and long time windows at the stapes and the ear canal. Perturbations were applied in $\delta(x)$. 238

Figure E.1.a-b: Illustration of the basin of entrainment in a Van der Pol Oscillator 243

List of Tables

Table 2.1: Model parameters for Neely and Kim's (1986) cat cochlea, and the revised quantities for the human cochlea.....25

Table A.1: Physical characteristics of the ear canal.....184
Table A.2: List of middle ear impedances and physical quantities being modelled.....188
Table A.3: Parameters of the middle ear model in cgs units. 189
Table A.4: Parameters of the middle ear model in SI units. 189
Table A.5: Lumped element parameters of the middle ear boundary.....199

List of Accompanying Material

A compact disc is included with this doctoral thesis that contains the following files:

- a PDF of this dissertation;
- high-resolution electronic versions of all figures used in this thesis, organised by chapter;
- MATLAB scripts and functions that generated the numerical results presented here;
    - a reference document called '`Guide to MATLAB Code.doc`' is included.
- animations of BM motion in two time domain simulations:
    - `Step_unstable_NLTD_simulation.wmv` – nonlinear time domain simulation of an unstable cochlea, given a step change in γ.
    - `Random_unstable_NLTD_simulation.wmv` – nonlinear time domain simulation of an unstable cochlea, given random changes in γ.

Declaration of Authorship

I, Emery M. Ku, declare that the thesis entitled *Modelling the Human Cochlea* and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- parts of this work have been published as:
    - Ku, E.M., Elliott, S.J. and Lineton, B. (2008). 'Statistics of instabilities in a state space model of the human cochlea,' J. Acoust. Soc. Am.; 124: 1068-1079.
    - Elliott, S.J., Ku, E.M., and Lineton, B. (2007). 'A state space model for cochlear mechanics,' J. Acoust. Soc. Am. 122: 2759-2771.

Signed: …….

Date:…….

Acknowledgements

I would like to thank Steve Elliott and Ben Lineton for their guidance, insights and expertise. In particular, Professor Elliott's unwavering support and passion for research have meant a great deal to me.

I would not be who or where I am today without my parents. My mother taught me to strive for excellence in all endeavours, and my father has always been my role model.

It has been my good fortune to share an office and several homes with a wide range of generous and welcoming individuals. I am indebted to them for their support and many kindnesses through the years.

Participating in two student societies, the Southampton University Hillwalking Club and the choir, Jazzmanix, has enriched my experience in the United Kingdom. I am grateful for the lifelong friends that I have made through these groups.

Finally, it is the US-UK Fulbright Commission that enabled me to come and study in England. I am proud to be associated with this institution, and hope that my time here has furthered Senator Fulbright's vision for deeper cultural understanding between nations.

List of Symbols

General Symbols

$\lvert\ \rvert$	absolute value of a quantity
$\angle$	phase of a quantity [cycles]
$\otimes$	convolution operator
Σ	summation operator
∂	partial differential operator
$\det(\ )$	determinant of a matrix
$\Re(\ )$	real part of a complex variable
$\Im(\ )$	imaginary part of a complex variable
$\mathbf{A}$	state space system matrix
A_{st}	area of the stapes footplate [m^2]
a_{st}	stapes acceleration [$m \times s^{-2}$]
$\mathbf{B}$	state space input matrix
b	ratio of ξ_b to ξ_p
$\mathbf{C}$	state space output matrix
c_1, c_2, c_3, c_4	damping terms in the micromechanical models [$N \times s \times m^{-3}$]
c_m, c_H	damping terms in the middle ear and helicotrema boundary elements [$N \times s \times m^{-3}$]
c_{TW}	TW wavespeed [$m \times s^{-1}$]
$\mathbf{D}$	state space feed-through matrix
$\mathbf{F}$	fluid-coupling matrix
f	geometric mean of two frequencies [kHz]
g	BM to IHC lever gain
H	height [m]
$h(t)$	impulse response of a transfer function

$\mathbf{I}$	unit matrix
j	$\sqrt{-1}$
k_1, k_2, k_3, k_4	stiffness terms of the micromechanical model [N×m^{-3}]
k_m	stiffness terms of the middle ear boundary element [N×m^{-3}]
L	length of the cochlea [m]
l	cochlear length scale [m]
$\mathbf{M}$	admittance matrix
m_1, m_2	mass of the BM and the TM in the micromechanical model [kg×m^{-2}]
m_m, m_h	mass of the middle ear and helicotrema boundary elements [kg×m^{-2}]
N	number of cochlear elements
$\mathbf{p}$	matrix of pressures
p_i	complex pole
p_a	active pressure [Pa]
p_d	pressure difference across the cochlear partition [Pa]
Q	volume velocity [m^3×s^{-1}]
$\int Q$	volume displacement [m^3]
$\mathbf{q}$	vector-matrix of source terms in the state space model
s	complex frequency variable, equal to $j\omega$ unless otherwise noted
t	time [s]
$\mathbf{U}$	fluid-coupled matrix of source terms to the state space model
$\mathbf{u}$	state space model input matrix
$\mathbf{x}$	state space state matrix
x	longitudinal position [m]
Y_{cp}	cochlear partition admittance [N^{-1}×s^{-1}×m^3]
$Z_1, Z_2, Z_3, Z_4,$	micromechanical impedance terms [N×s×m^{-3}]
Z_{cp}	cochlear partition impedance [N×s×m^{-3}]
Z_{in}	acoustic impedance looking in to the cochlea [N×s×m^{-5}]
Z_{out}	acoustic impedance looking out of the cochlea [N×s×m^{-5}]
z_i	complex zero
Δ	thickness of a cross-sectional element of the cochlear model [m]
Δf	difference between adjacent unstable frequencies [kHz]
γ	micromechanical feedback gain
δ	saturation point of the Boltzmann function

ζ	damping ratio
κ_{TW}	complex TW wavenumber [mm^{-1}]
λ_{TW}, λ_{peak}	TW wavelength, wavelength of the TW at its peak [mm]
ξ_b	BM displacement [m]
ξ_c	relative displacement between TM and BM [m]
ξ_H	displacement at the helicotrema boundary [m]
ξ_m	displacement of the middle ear boundary [m]
ξ_p	average BM displacement over the radial cross-section [m]
ξ_t	TM displacement [m]
ρ	cochlear fluid density [$kg \times m^{-3}$]
ς	undamping ratio
σ	real part of a pole [ms^{-1}]
τ_{TW}	cochlear group delay [ms]
ω	angular frequency [$radians \times s^{-1}$]

Specific Symbols used in Chapters 3-4

$\mathbf{A_E}$, $\mathbf{B_E}$, $\mathbf{C_E}$, $\mathbf{D_E}$	elemental matrices in the state space model of the uncoupled cochlea
k_{rw}	stiffness of the round window [$N \times m^{-3}$]
$[x_1, x_2, x_3, x_4]_n$	states of the n^{th} state space element in the cochlea
λ_{sin}	wavelength of a sinusoidal distribution of feedback gain [mm]
ϕ	phase [radians]

Specific Symbols used in Appendix A and E

A	area of the ear canal [m^2]
d	diameter of the ear canal model [m]
C	mechanical compliance [$Pa^{-1} \times m^3$]
c_0	speed of sound in air [$m \times s^{-1}$]
k_1, k_2, k_3,	middle ear transformer ratios
k_{st}	stiffness of the middle ear model [$N \times m^{-3}$]
G_{ME}^{RT}	round-trip gain through the middle ear
L	inertia [$Pa \times s^2 \times m^{-3}$]
M1, M2	forward- and backward- pressure gains through the middle ear model
P_{ec}	pressure in the ear canal of a two-port network [Pa]
P_{ed}	pressure at the ear drum of a two-port network [Pa]
P_{st}	pressure at the stapes of a two-port network [Pa]
Q_{ec}, Q_{st}	volume velocity at the ear canal and stapes [$m^3 \times s^{-1}$]
R	mechanical resistance [$Pa \times s \times m^{-3}$]
$\mathbf{T}$	two-port network transfer matrix
Z_1, Z_2, Z_3, Z_4,	grouped impedances in the middle ear model [$Pa \times s \times m^{-3}$]
Z_{in}, Z_{out}	input- and output- impedances of the two-port network [$Pa \times s \times m^{-3}$]
γ	damping factor in the Van der Pol equation
ρ_0	density of air [$kg \times m^{-3}$]
ω_n, ω_d,	natural and driven frequencies in the Van der Pol equation [$radians \times s^{-1}$]

List of Abbreviations

1-D	one-dimensional
3-D	three-dimensional
BM	basilar membrane
CEOAE	click-evoked otoacoustic emission
CF	characteristic frequency
CP	cochlear partition
CWT	continuous wavelet transform
DD	directly-determined
DFT	discrete Fourier transform
DNL	derived nonlinear
ESLE	extremely short latency emission
IHC	inner hair cell
OAE	otoacoustic emission
ODE	ordinary differential equation
OHC	outer hair cell
PMD	preferred minimum distance
RMS	root-mean-square
SOAE	spontaneous otoacoustic emission
SPL	sound pressure level
TBOAE	tone-burst otoacoustic emission
TEOAE	transient-evoked otoacoustic emission
TM	tectorial membrane
TW	travelling wave
WKB	Wentzel-Kramers-Brillouin

General Comments

- The term ‘measurement’ is used in this thesis to refer to any experimental measurements performed on live or dead subjects, whereas the term ‘simulation’ refers to any results generated by numerical simulation.
- The term ‘baseline’ is used to describe a cochlear model that is active, passive, linear or nonlinear, but does not contain any deviations from standard, uniformly varying parameters; if a model contains deviations from normal parameter values, the term ‘perturbed’ is applied.
- Matrices are presented in upper case, bold typeface; vectors are presented in lower case, bold typeface.
- $\dot{}$ and $\ddot{}$ indicate the first and second derivatives of a variable with respect to time.
- The notation of a given variable’s dependency on x, ω or t is occasionally suppressed for convenience.
- Quantities are often expressed in millimetres and milliseconds as these units are generally better-suited to the spatial- and temporal- scales of the cochlea and its activity.

Chapter 1

Introduction

1.1 The Human Cochlea

Of the various biological systems under study, the human cochlea is a particularly fascinating and challenging organ to model. The cochlea represents the last mechanical stage of hearing. It has great acoustical resolving power in time, frequency and intensity, and exhibits a wealth of nonlinear phenomena. Many aspects of its function are still not well-understood, even though it has been examined for hundreds of years (e.g. Helmholtz, 1874); this is primarily due to the difficulties associated with direct mechanical measurements of the cochlea. For instance, its dimensions and motion are on extremely small magnitudes. This is further complicated by its physical inaccessibility, being located deep within the temporal bone.

The goal of this doctoral thesis is to add to the understanding of how the human cochlea operates by performing computer simulations of its dynamics. What follows is an introduction to the biology and observed features of the mammalian cochlea. Also included is an overview of the sounds that can propagate out of the cochlea, known as otoacoustic emissions (OAEs). OAEs represent an important epiphenomenon of the cochlea as their analysis can reveal clues to cochlear function without requiring invasive surgical techniques.

1.1.1 Biology and measurements

The cochlea is a fluid-filled organ located in the inner ear. Its primary function is to perform mechanical pre-processing of incoming acoustical pressure waves and to convert this energy into neural impulses that are sent to the auditory cortex. In addition to amplifying input signals, the cochlea maps the motion induced by the various frequency components of a sound to different positions along the cochlea. This spatial filtering allows a large number of neuronal pathways, each with a limited bandwidth, to relay much of the information contained within the acoustical signal to the brain. Figure 1.1 shows the cochlea's location relative to the outer and middle ears.

The middle ear bones, the malleus, the incus and the stapes, perform an impedance matching between the air in the outer ear and the fluid in the cochlea; this is achieved through the reduction in surface area between the eardrum and the stapes footplate, in addition to the lever arm that arises due to the geometric arrangement of these bones. A more in-depth discussion of the middle and outer ears can be found in Appendix A.

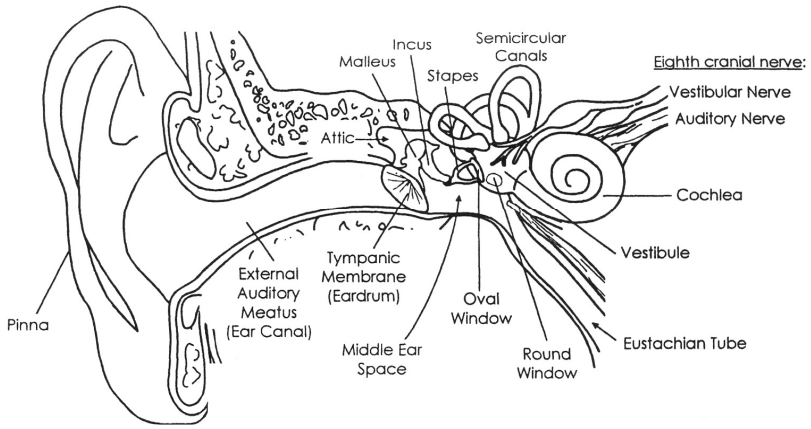


Figure 1.1: The cochlea in relation to the outer and middle ears. Reproduced from 'Hearing: an introduction to psychological and physiological acoustics,' by S.A. Gelfand, Copyright (1998), with permission from Marcel Dekker.

Sound waves incident upon the eardrum induce motion in the middle ear bones, which in turn cause the stapes footplate to produce pressure waves in the cochlear fluid. Figure 1.2 shows a simplified view of the scalae, or chambers, of an uncoiled cochlea. The inward displacement of the stapes at the oval window results in a near-instantaneous outward displacement of equivalent volume at the round window due to the incompressibility of the cochlear fluid. This equalizes the overall pressure in the scalae

(Pickles, 2003). However, a slower travelling wave (TW) is observed on the basilar membrane (BM), a thin sheet of material in the cochlea. This TW propagates from base to apex following a stimulus (von Békésy, 1949). Figure 1.3 shows the instantaneous BM and fluid motion associated with the TW in the scalae produced by a tonal stimulus.

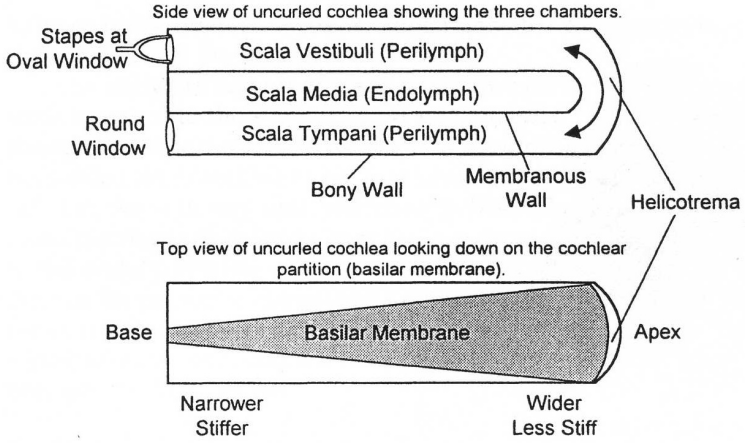


Figure 1.2: Schematic representation of the uncoiled cochlea. Reproduced from 'Hearing: an introduction to psychological and physiological acoustics,' by S.A. Gelfand, Copyright (1998), with permission from Marcel Dekker.

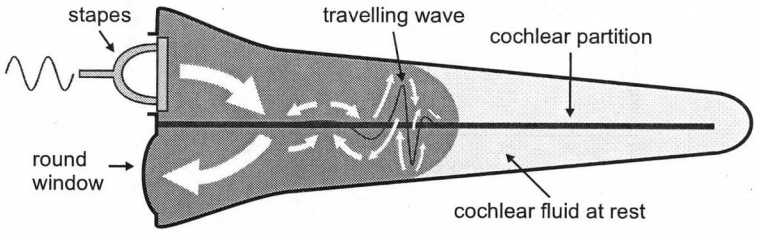


Figure 1.3: Schematized illustration of the BM travelling wave and fluid flow given sinusoidal excitation at the stapes. Redrawn after Trends in Neurosciences, 21, Nobili, R., Mammano, F. and Ashmore, J., 'How well do we understand the cochlea?' 159-167, Copyright (1998), with permission from Elsevier.

The BM is stiff and narrow at the base of the cochlea and broadens to become wider and floppier at the apex, as schematized in Figure 1.2. This variation in the BM's mechanical properties results in a natural or 'passive' tuning of the response of the cochlea; higher frequencies resonate near the base, and lower frequencies near the apex. At a given excitation frequency, speed of TW and its local wavelength decrease as it approaches its peak. This effect is similar to the behaviour of ocean waves which get taller and narrower as they encounter shallower waters. Mechanically speaking, a TW generated by a

sinusoidal excitation travels quickly in the basal, stiffness-dominated region. It then slows and reaches a peak at its resonant point where the BM impedance is lowest; this location is defined as the 'characteristic place.' Apical of the characteristic place, the impedance of the BM is mass-dominated and the TW is extinguished (von Békésy, 1949). This spatial mapping of tones, sometimes referred to as 'tonotopy,' is further enhanced by active elements located in the organ of Corti (OC), a set of specialized cells that sits on the BM. Figure 1.4 shows a cross-section of the cochlea and the three scalae, while Figure 1.5 presents a detailed view of the OC.

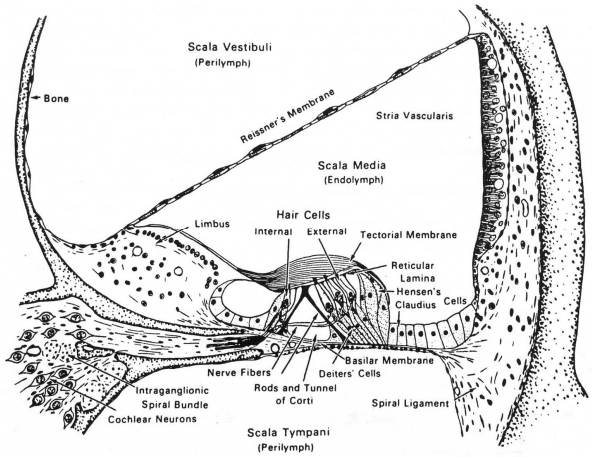


Figure 1.4: Cross-section of a single turn of the cochlea. Reproduced from 'Hearing: an introduction to psychological and physiological acoustics,' by S.A. Gelfand, Copyright (1998), with permission from Marcel Dekker.

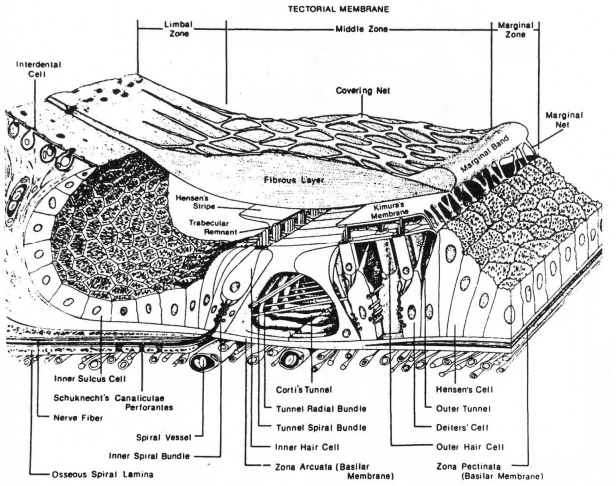


Figure 1.5: Detail view of the organ of Corti. Reprinted from Hearing Research, 22, Lim, D.J., 'Functional structure of the organ of Corti: a review,' 117-146, Copyright (1986), with permission from Elsevier.

The OC is critical to the function of the cochlea. The travelling wave on the BM propagates through the cochlea as a result of a local pressure differences above and below the cochlear partition (CP). The bending of the BM results in a shearing motion against a gelatinous flap that sits above it, the tectorial membrane (TM). Embedded in the TM are hair-like cells known as stereocilia that protrude out of the outer hair cells (OHCs). The deflection of the stereocilia tip links mechanically opens and closes ion channels which are believed to be associated with the transduction process. This in turn induces the electro-chemical expansion and contraction of the OHCs in a phenomenon known as somatic motility (Ashmore, 1987). The forces generated by the OHCs are believed to actively amplify the TW and sharpen its pattern of activity along the BM (Sellick *et al*., 1982). Of course, such a cochlear amplifier (CA) requires a source of energy. Adjacent to the scala media is the stria vascularis, which produces the positively charged fluid that fills this chamber. This liquid is known as endolymph, and it exhibits a positive resting potential on the order of +100 mV relative to the other chambers (Gelfand, 1998).

Although all research to date regarding the mechanics of the OHCs has been performed *in vitro*, the electromotility of these cells has been demonstrated in a number of experiments (Brownell *et al*., 1985). Figure 1.6 shows the change in intracellular voltage given variations in pressure for both an inner hair cell (IHC) and an OHC.

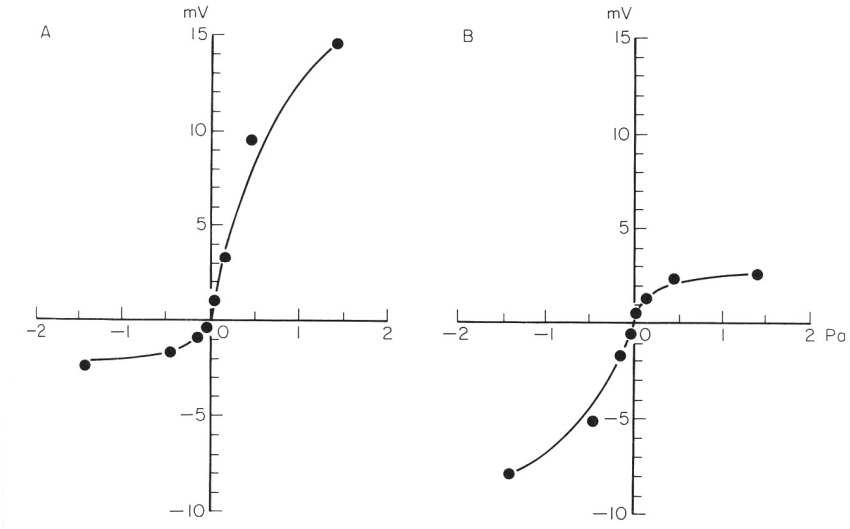


Figure 1.6: Input-output functions of an IHC (left) and OHC (right) from the base of a guinea pig cochlea. Intracellular voltage (vertical axis) is plot against stimulus pressure (horizontal axis). Reprinted from the Journal of Physiology, 383, Cody, A.R. and Russell, I.J., 'The response of hair cells in the basal turn of the guinea pig cochlea to tones,' 551-569, Copyright (1987), with permission from the Physiological Society.

It is clear that the transduction characteristics of inner- and outer-hair cells are both non-symmetrical and saturating, and thus nonlinear. Researchers have fitted these responses to Boltzmann functions with good agreement in the past (e.g. Kros *et al*., 1992). Note that in later figures, the response of the Boltzmann function has the shape of the IHC transduction measurement, rather than the inverted shape for the OHC response shown in Figure 1.6.

Figure 1.7 shows the exposed stereocilia on the upper surface of a chinchilla's organ of Corti, with the TM removed. In contrast to OHCs which are believed to amplify the TW motion, the IHCs are understood to act as sensors which detect motion and encode this information via the release of neurotransmitters. There are approximately 30,000 sensory neurones which carry this information to the central nervous system (Pickles, 2003). However, there are also a much smaller number of 'efferent' nerve fibres which convey instructions from the brain to the base of the hair cells, particularly the OHCs. The exact function of cochlear efferents is still unknown, though they are likely to act as an adaptive control signal pathway.

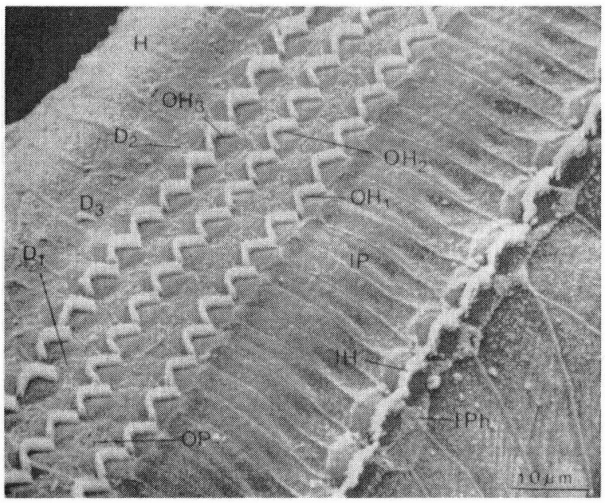


Figure 1.7: The upper surface of a chinchilla organ of Corti. Stereocilia of inner and outer hair cells are labelled as IH and OH. Also marked are Hensen's cells (H), Deiters' cells (D), outer and inner pillar cells (OP, IP), and inner phalangeal cells (IPh). Reprinted from Hearing Research, 22, Lim, D.J., 'Functional structure of the organ of Corti: a review,' 117-146, Copyright (1986), with permission from Elsevier.

The importance of the organ of Corti is underscored when it is physiologically compromised. Cochlear damage may arise due to a variety of factors, such as acoustical overstimulation or exposure to ototoxic substances. The presence of cochlear injury may

in turn be manifest as a reduction in hearing sensitivity or frequency discrimination. From an engineering standpoint, the resolving capabilities of a healthy human cochlea are remarkable.

The cochlea is able to discern minute changes in intensity, time and frequency. For instance, the average human can detect a variation in frequency of less than 1% across its bandwidth, which spans approximately three decades (Gelfand, 1998). One of the most salient features of the cochlea is the dynamic range it possesses; this is approximately 140 dB SPL—the loudest bearable sound is 10,000,000 times greater in pressure amplitude compared to the softest detectable sounds (Gelfand, 1998). At the lower threshold of human hearing, the cochlea can detect motion that is smaller than the width of a hydrogen atom (Rhode, 1984). At the upper end of hearing intensities are sounds that are comparable to those generated by a rifle fired at close range.

The staggering dynamic range of the cochlea is in part due to the active process within the cochlea, termed the cochlear amplifier (CA). At low amplitudes of motion, the CA behaves linearly and provides approximately 45 dB of amplification (Pickles, 2003). If not for the contribution of the OHCs, sounds on the order of 0 dB SPL would be imperceptible. As stimulus levels increase, the CA begins to saturate. The *relative* contribution of the OHCs to the motion of the BM begins to decrease, and thus the *effective* gain is also reduced. The magnitude of the BM response grows less than linearly within this saturating range; this compressive behaviour is observed for excitation levels of approximately 40-80 dB SPL presented at the outer ear (Pickles, 2003). Finally, at the highest levels of tolerable pressure levels, beyond approximately 90 dB SPL, the growth of the BM motion again becomes approximately linear as the active contributions of the OHCs are negligible. The variation of the growth of BM motion for increasing stimulus levels from linear → compressive → linear is illustrated in Figure 1.8.

Figure 1.9 illustrates the measured frequency response of the BM displacement at a single point along the cochlea as a function of the external sound pressure level. The response is sharply tuned at low levels and broadens with increasing amplitude. The frequency of maximal response also appears to decrease with increasing amplitudes. This is often referred to as the ‘half-octave shift’ in the literature, where the frequency of the maximum BM vibration shifts downward by approximately half an octave with increasing driving levels (Johnstone *et al.*, 1986).

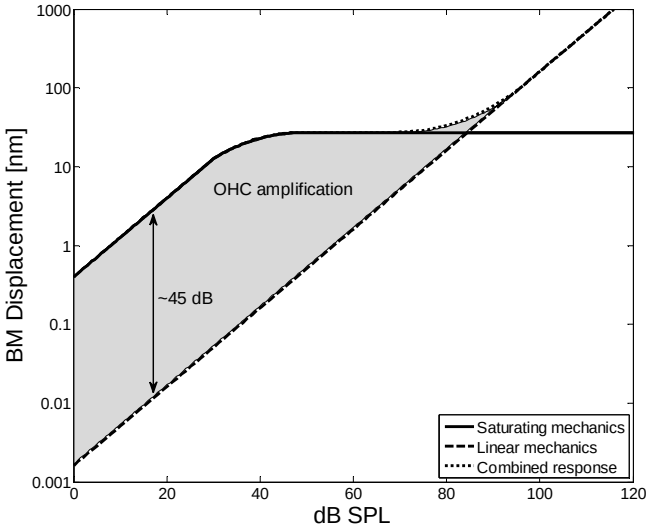


Figure 1.8: Nonlinear growth and compression of the BM response. Redrawn after Pickles, J.O., 'An Introduction to the Physiology of Hearing,' Second Edition, Copyright (1988), with permission from Elsevier.

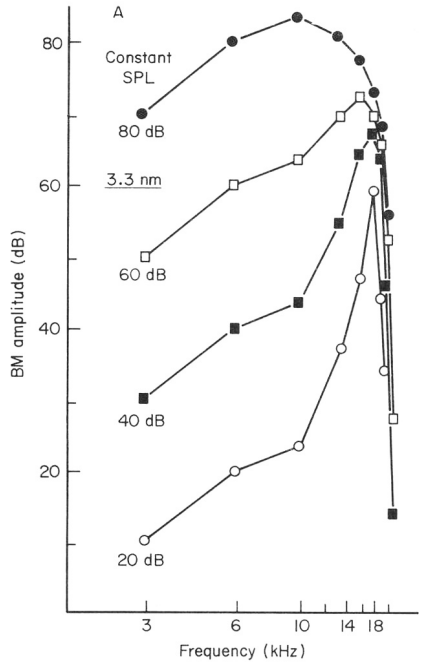


Figure 1.9: Variation in the amplitude of the BM response with level. Reprinted from Pickles, J.O., 'An Introduction to the Physiology of Hearing,' Second Edition, Copyright (1988), with permission from Elsevier.

It is important to note that many of the above observations and measurements are restricted to responses in the basal half of the mammalian cochlea. This is due to the

inaccessibility of the middle turns. Recent research suggests that the mechanics and active components of the apex may operate in a manner that is fundamentally different to that in the base (e.g. Guinan *et al.*, 2005).

1.1.2 the Dominant source of cochlear nonlinearity

There are many potential sources of nonlinearity in the cochlea, ranging from the complex mechanical and electrochemical process of OHC stimulation and activation to the variation of BM stiffness with driving level (Patuzzi, 1996). What is of greatest relevance to the cochlear modeller, however, is which of these nonlinearities dominates the global system response. As we have seen, the most fundamental nonlinear effect in the cochlea is the compressive growth of the BM motion at moderate levels. This is not an unintentional consequence of cochlear mechanics; this compression allows the IHCs to process a much wider range of sounds than would otherwise be possible. As it is the physiological limitations of the amplification provided by OHCs that give rise to this effect, the saturation of this feedback force can be considered the primary source of nonlinearity in the cochlea.

The nonlinearity of the CA gives rise to an abundance of nonlinear phenomena in the cochlea. For instance, harmonic distortion in the BM response to tones has been measured in vivo by experimentalists (e.g. Cooper, 1998). In addition, the simultaneous application of two tones can produce a variety of interactions, such as the generation of new tones (intermodulation distortion products), and suppression or enhancement of the response of one of the applied frequencies (Robles and Ruggero, 2001). Some researchers have attempted to model these features with the inclusion of a saturating active feedback force (e.g. Kanis and de Boer, 1994). This effectively limits the amplitude of any motion generated by the OHCs.

1.2 Otoacoustic Emissions

The existence of an active mechanism in the cochlea was first hypothesized by Thomas Gold in 1948. He believed that an electromechanical action is necessary to counteract the heavy viscous damping in the fluid-filled cochlea. Gold (1948) went on to propose that a perturbation may ‘bring an [active] element into the region of self-oscillation, when it is normally so close to [instability].’ Although the concept of a CA is largely taken for

granted today, it met a great deal of scepticism when first proposed. The field of cochlear mechanics in the 1940s was very much dominated by the research of Georg von Békésy, the scientist who discovered the existence of a tonally-generated travelling wave in the cochleae of human cadavers (von Békésy, 1949; Hall, 2000). As von Békésy's experimental work dealt with preparations of dead cochleae, it is unsurprising that his conclusions regarding the mechanics of the organ were 'passive' in nature. It would be another three decades until Gold's ideas were re-examined. In 1978, David Kemp published findings of sounds measured in the ear canal that had a cochlear origin (Kemp, 1978); thus, the field of otoacoustic emission (OAE) research was born.

A generally used definition of an OAE is any sound that is generated from within the cochlea and externally measured (Hall, 2000). For the purposes of this work, an OAE is defined as a variation in the pressure in the ear canal, or at the base of the cochlea, that was generated in the cochlea. It is traditional in the literature to classify the emission type by the stimulus. For instance, the self-oscillating emission that was predicted by Gold (1948) is now referred to as a spontaneous otoacoustic emission (SOAE), whereas emissions that are generated by a short click stimulus are termed click-evoked otoacoustic emissions (CEOAEs). This predominant system of nomenclature is adopted in this thesis for clarity and in order to directly compare model results with clinically-measured data.

During the last ten years, there has been a move within the literature toward a consensus regarding the underlying causes of OAE generation. Shera and Guinan (1999) proposed that OAEs arise due to both linear reflection and nonlinear distortion. Cochlear reflections are believed to arise as a result of small imperfections at fixed locations along the BM, whereas distortion is thought to be a by-product of the (frequency-dependent, place-shifting) nonlinear amplification process. These two very different mechanisms have been contrasted as 'place-fixed' or 'wave-fixed' in the literature (Kemp, 1986). It is now generally accepted that all forms of evoked emissions are a combination of both linear and nonlinear mechanisms, though the dominance of each generation mechanism in various circumstances is still being debated.

There is also evidence to suggest that all forms of OAEs are related and directly tied to the sensitivity of hearing (Zwicker and Schloth, 1984; McFadden and Mishra, 1993; Talmadge and Tubis, 1998; Shera and Guinan, 1999). Indeed, physiological insult can reduce or remove the presence of all forms of OAEs. For this reason, evoked OAEs have

been investigated extensively for their clinical applications. At the time of this writing, both CEOAEs and distortion-product OAEs (DPOAEs) screenings are commonly used to detect hearing defects (Hall, 2000). Neonates and other children who are too young to cooperate in conventional hearing tests are tested for CEOAEs in many Western countries. All forms of emissions are averaged in time in order to minimize any physiological or environmental noise contamination present in the signal.

What follows is a brief overview of the primary classes of OAEs; detailed discussions of SOAEs and CEOAEs are reserved for the relevant chapters.

1.2.1 SOAEs

SOAEs are defined as low-amplitude, narrowband sounds that are emitted from the cochlea without any stimuli. They are believed to be a feature of a normally functioning CA, as they are commonly found in an estimated range of 33% to 70% of all normally-hearing ears (Penner and Zhang, 1997; Talmadge *et al*., 1993). Furthermore, multiple emissions (as many as 35) are common in individuals with SOAEs (Probst et al., 1991). The time-averaged spectrum of a sample SOAE measurement is given in Figure 1.10.

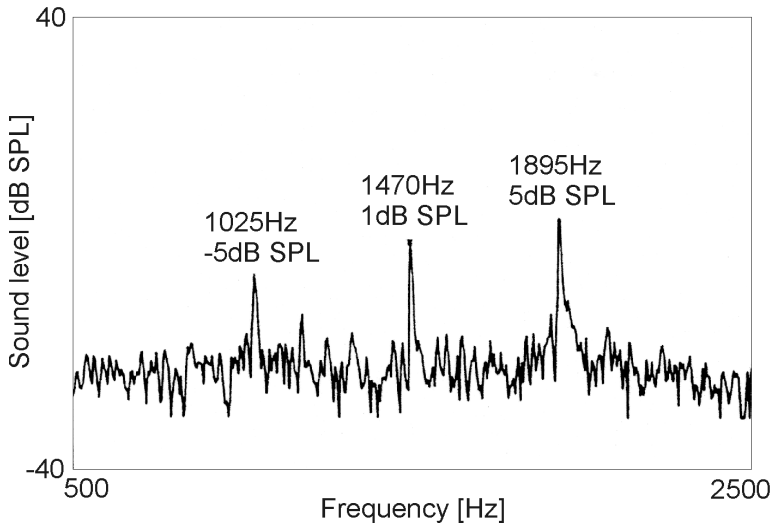


Figure 1.10: Example of a SOAE measurement in a human. Adapted with permission from the Journal of the Acoustical Society of America, 89, Probst, R., Lonsbury-Martin, B.L., and Martin, G.K., 'A review of otoacoustic emissions,' 2027-2067, Copyright (1991).

Though advances in detection techniques have shown a clear rise in the measured incidence of such emissions, the absence of SOAEs does not necessarily imply cochlear dysfunction (Hall, 2000). The amplitude of spontaneous emissions are most commonly

observed between -5 and 15 dB SPL in the ear canal, though unusual cases have been reported with emissions as loud as 55 dB SPL (Probst *et al.*, 1991; Hall, 2000). Spontaneous emissions occur in a wide range of frequencies, though most are detected in the 1 – 2 kHz band. SOAE frequencies of up to 7 kHz are regularly found, though many studies do not measure lower or higher frequencies due to equipment and noise floor limitations.

Whereas the amplitude of an emission may vary with time, its frequency often remains constant (within 1%) for years at a time (Hall, 2000). Another notable characteristic of SOAEs is that there is a regularly-observed (log-normalised) average spacing between adjacent SOAE frequencies (Dallmayr, 1985, 1986; Talmadge et al., 1993; Braun, 1997). The implications of this attribute for understanding cochlear mechanics are discussed in Chapter 4.

1.2.2 CEOAEs

Click-evoked emissions represent the cochlea's response to a short-duration, wide-band stimulus. They are present in approximately 98% of normally hearing adults, and thus are well-suited to clinical applications (Probst *et al.*, 1991). The magnitude of the emission is much lower than the stimulus level and it is delayed as well; for these reasons, CEOAEs and other transiently-evoked OAEs (TEOAEs) are sometimes colloquially referred to as '*Kemp's echoes*' (Hall, 2000).

Although the electrical stimulus sent to a transducer is a rectangular pulse of typically 100 μs width, there is residual ringing within the ear canal that lasts for 3-5 ms. The duration of this transient can be exarcerbated if the OAE probe is poorly fitted (Hall, 2000; Harte, 2004). This has been attributed to both the response of the transducer and the transient response of the middle and outer ears. Various techniques have been developed to remove this stimulus artefact, as discussed in Chapter 5.

Figure 1.11 shows the CEOAE response of a healthy young adult as displayed by commercially available equipment and software. Note that the first four ms which include the stimulus are displayed in a separate panel (upper-left) whereas the largest window shows only the longer-latency result (the OAE); the difference between the scales of these two responses is on the order of a factor of 1000.

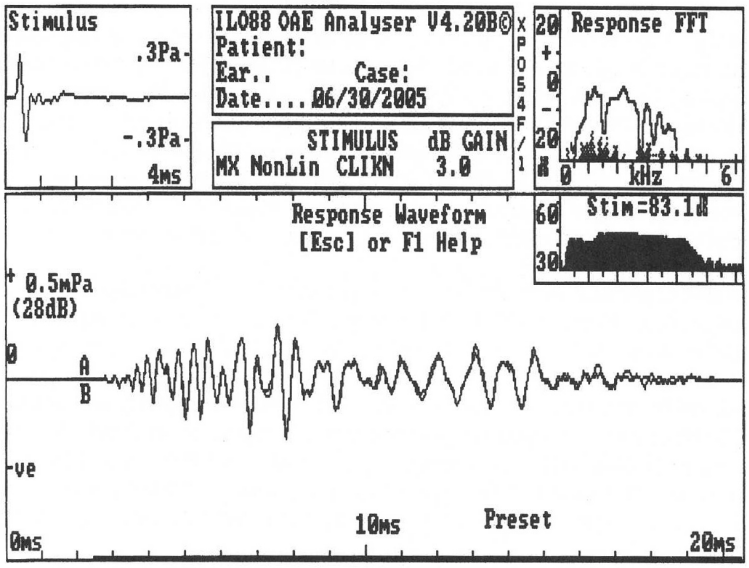


Figure 1.11: Example of a typical CEOAE measurement in a normal-hearing human subject obtained with commercially available equipment. Reprinted from Robinette, M.S., and Glattke, T.J., 'Otoacoustic Emissions: Clinical Applications,' Third Edition, Copyright (2007), with permission from Thieme Medical Publishers.

The instantaneous frequency of the CEOAE shows considerable variation as a function of time, with frequency varying inversely with latency. This is believed to be due to the connection between the dispersive characteristics of the cochlea and its tonotopic mapping (Greenwood, 1990; Sisto and Moleti, 2002). It has also been noted that the frequency spectrum of the CEOAE is strongly dependent on that of the click stimulus itself (Zwicker and Schloth, 1984; etc). As shown in the upper-right panel of Figure 1.11, the spectrum of the emission is concentrated in the 0.5-2 kHz region; this is a typical feature of many OAEs and is believed to be due to the band pass-like transmission characteristics of the middle ear (Kemp and Chum, 1980; Puria, 2003; Ku *et al.*, 2008).

As with other forms of evoked emissions, the amplitudes of CEOAEs grow proportionally at low levels (below 20-30 dB SPL) and saturate heavily beyond this point (Kemp, 1979; Zwicker and Schloth, 1984; Probst *et al.*, 1986; Probst *et al.*, 1991). The growth of CEOAE amplitudes above and beyond the saturating range of stimulus levels is rarely reported, as prolonged exposure at these levels can cause permanent cochlear damage. In addition, physiological protection mechanisms may introduce noise or otherwise skew the recorded signal (Hall, 2000). The latency of various frequency

components of the CEOAE also appears to vary inversely with stimulus level (Sisto and Moleti, 2007). Representative simulations of CEOAEs are presented in Chapter 5.

1.2.3 Otherwise-evoked OAEs

There are several other classes of evoked OAEs that are commonly discussed in the literature: tone-burst OAEs (TBOAEs), stimulus-frequency OAEs (SFOAEs), and distortion product OAEs (DPOAEs). TBOAEs are similar to CEOAEs given their transient nature, though CEOAEs are typically preferred clinically as they elicit a broader cochlear response. SFOAEs are most apparent at low levels of stimulation (e.g. Zwicker, 1990). However, SFOAEs are of limited clinical use as the time required to test for their existence is much longer than other forms of OAEs (Hall, 2000).

Distortion-product OAEs are generated when two (or more) tonal stimuli are presented that are nearby in frequency (Probst *et al.*, 1991). The nonlinear aspects of cochlear mechanics result in the mechanical intermodulation of the two 'primary' tones which can produce numerous 'secondary' frequencies. The frequencies of the DPOAEs depend on the relative spacing of the primaries, the strongest of which occur at $2f_1 - f_2$ in humans, where f_1 is the lower frequency. Whereas SOAEs and TEOAEs are believed to be primarily due to 'place-fixed' reflection mechanisms, DPOAEs at moderate levels are more likely to depend on 'wave-fixed' distortion (Shera and Guinan, 1999).

1.3 Models of the Cochlea

Historically, the formulation of cochlear models has been driven by experimental measurements. For instance, von Békésy's measurements of travelling waves in dead cochleae inspired passive models for many years. As neural tuning curves showed much sharper responses than were present in the published TW findings of the day, researchers sought a neurological 'second filter' to explain the differences (Hubbard and Mountain, 1996). It was not until Kemp's discovery of OAEs decades later coupled with measurements of BM motion *in vivo* that researchers (fervently!) began to incorporate active elements into their models. With recent advances in computer processing capabilities, scientists now have the tools to simulate the behaviour of complex, nonlinear systems that lack closed-form analytical solutions. These numerical results may in turn

shed light on the intricacies of cochlear mechanics and thereby suggest further informative experiments.

This subsection gives a short overview of the wide variety of cochlear models that have been devised. Special attention is given to lumped-element models, as the system applied in this thesis falls in that category.

Soon after the findings of nonlinear BM activity and the existence of OAEs were published (Rhode, 1971; Rhode, 1974; Kemp, 1979), the first active cochlear models began to appear (e.g. Kim *et al.*, 1980). Initial mechanical models were formulated in the frequency domain and assumed active un-damping at a site basal to the characteristic place, but included only a single degree of freedom to represent the dynamics of the CP. This fixed the spatial distribution of undamping in the model, and thus the pattern of impedances was only valid for one frequency (Hubbard and Mountain, 1996). Later formulations included a second degree of freedom to represent the TM above the BM (Zwislocki and Kletsky, 1979; Allen, 1980; Neely and Kim, 1983, 1986). This allowed the active response to be generalized over the entire range of locations along the CP, and thus the entire spectrum of audible frequencies (Neely and Kim, 2007).

Early time domain simulations of cochlear models were used to demonstrate the stability of active models and to begin to incorporate nonlinearities, though computational limitations were restrictive (e.g. Diependaal *et al.*, 1987). Some later work in linear frequency-domain modelling attempted to determine the impedance of the BM by 'inversely' analysing experimental measurements (Zweig, 1991; de Boer, 1995). The model of Zweig (1991) is a departure from most early models in that it assumes the BM motion is like a negatively damped harmonic oscillator that is stabilised by a time-delayed negative feedback force. Yet others sought to refine model predictions by adding more degrees of freedom to the CP (Parthasarathi *et al.*, 2000), expanding solutions into multiple spatial dimensions (Kolston, 1999), or including so-called 'feed-forward' or 'feed-backward' coupling between adjacent BM impedances (Geisler and Sang, 1995; Fukazawa and Tanaka, 1996; Xin *et al.*, 2003).

As an alternative to traditional active elements, nonlinear limit cycle generators (such as the Van der Pol oscillator) are also referenced in the literature (Van Netten and Duifhuis, 1983; Duifhuis *et al.*, 1985; Wit, 1986; van Hengel *et al.*, 1996). These are often applied for their nonlinear characteristics which bear a striking resemblance to those of

SOAEs. Still, another subset of researchers regards the TW as an epiphenomenon and believes local resonance and or 'fast' compression waves in the fluid to be dominant in the cochlea (Bell, 2001).

This bewilderingly long list of only the most-oft-referenced models in the literature betrays the many varied (and often contradictory) views of how the cochlea functions. However, it is important to remember that the validity of any model is always restricted to a particular set of conditions. One of the goals of this exploration is to simulate OAEs in humans given a mechanical representation of the cochlea. As such, a relatively simple model that demonstrated a number of key features of the cochlea was selected as the starting point for the work presented in this thesis.

1.3.1 the Neely and Kim (1986) model

The model of Neely and Kim (1986) is an active, lumped-element representation of a cat cochlea. It was published as a linear frequency-domain formulation, though its mechanical basis lends itself well to simulation in the time-domain. While its publication date is but a few years shy of the birth date of the author, the Neely and Kim (1986) model nevertheless exhibits a number of fundamental characteristics of the cochlea:

1. inclusion of an active element in the cochlear micromechanics that enhances the amplitude of the TW
2. tonotopic tuning that is sharp when active and broad when passive

These key features also allow for the simulation of OAEs when an appropriate middle ear representation is included. In addition, the model is based upon the structure of the biology; this lends the investigator the ability to directly simulate the effect of measured or inferred changes in physiology by applying modifications to the relevant mechanical parameters (or vice versa). For instance, the amplification provided by the active element is controlled by a scalar which is related to OHC function. However, there is still some degree of uncertainty regarding the effective mechanical properties of the cochlea; as such, the parameters used here are inferred from known characteristics of the TW. Research in this area is still ongoing (e.g. Newburg and Mountain, 2008).

Through the course of this research and deeper investigations of the literature, a number of contradictions between simulation results and experimental measurements have become apparent. These discrepancies are listed and discussed in Chapter 6.

1.3.2 Causes of variations in feedback gain

In a biological mammalian cochlea, the amplification generated by the OHCs can deviate from its nominal linear value due to a variety of circumstances, some of which are listed below:

a) Natural, uniform variations resulting in lower-than average auditory acuity
b) Static, place-fixed variations due to developmental randomness
c) Instantaneous, wave-fixed nonlinear saturation
d) Temporary or permanent cochlear pathology following noise damage, exposure to ototoxic substances, presbycusis (age-related hearing loss), etc.
e) Time-varying physiological factors
    i. Overstimulation recovery (Kemp and Brill, 2008)
    ii. Fatigue
    iii. Postural changes (de Kleine *et al*., 2000)
f) Externally applied voltages/currents

Items a-c) are addressed in this work, whereas d-f) are left to future generations of cochlear modellers.

1.4 Aims and Thesis Structure

One of the initial aims of this doctoral research was to simulate OAEs using an analytical, physiologically-based model of the human auditory system. It was hoped that these simulations would shed light on how the cochlea functions. Various investigations carried out here have shown that perturbations in the local feedback gain along the BM can have important consequences for the global performance of the cochlea. In order to better understand these effects, the mechanical parameters of the Neely and Kim (1986) model were updated for a human (Chapter 2) and the system was recast in a state space formulation (Chapter 3). The state space model provides a straightforward method for quickly and unambiguously analysing system stability, which is critical to understanding the formation of SOAEs (Chapter 4). In addition, the state space formulation is inherently based in the time domain, which allows for the simulation of transient phenomena such as CEOAEs (Chapter 5).

The inclusion of a saturation nonlinearity in the feedback loop allows for the meaningful simulation of unstable cochleae and the effect of increasing stimulus amplitudes. Nonlinear simulations of unstable cochleae are performed in Chapter 4, and

the amplitude-dependence of the cochlear response to tones and clicks is simulated in Chapters 3 and 5, respectively. The findings of this research are summarized and discussed in Chapter 6.

1.5 Contributions

The primary contributions of this work are as follows:

1. Refinement of a frequency-domain model of cochlear mechanics such that it now exhibits TW attributes pertinent to the generation of OAEs in humans, similar to experimentally measured and inferred values.
2. Co-development and implementation of a state space (time-domain) formulation of the refined cochlear model in MATLAB which includes a basal boundary condition based on physiological measurements.
3. Application of the state space model to validate one of the prevailing theories of SOAE generation.
4. Time domain simulations of cochlear motion and the complete process of evoking and measuring OAEs from the ear canal.
5. Comparison of time- and frequency-domain responses given static (linear) vs. dynamic (nonlinear) variations in feedback gain.

Some of these findings were shared with the wider academic community through journal papers and conference presentations:

- Ku, E.M., Elliott, S.J. and Lineton, B. (2008). 'Statistics of instabilities in a state space model of the cochlea,' J. Acoust. Soc. Am., 124, (2), 1068-1079.
    - this manuscript was selected for inclusion in the August 15, 2008 issue of the Virtual Journal of Biological Physics Research.
- Elliott, S.J., Ku, E.M. and Lineton, B. (2007). 'A state space model for cochlear mechanics,' J. Acoust. Soc. Am., 122, (5), 2759-2771.
- Ku, E.M., Elliott, S.J. and Lineton, B. (2008). 'Periodicity in the spectrum of modelled spontaneous otoacoustic emissions,' Proc. 10th Int. Workshop on the Mech. of Hearing, Keele University, U.K.

- Elliott, S.J., Ku, E.M. and Lineton, B. (2008). 'Time domain model of a nonlinear inhomogeneous cochlea,' Proc. 10th Int. Workshop on the Mech. of Hearing, Keele University, U.K.
- Ku, E.M., Elliott, S.J. and Lineton, B. (2008). 'Instabilities in a state space model of the human cochlea,' Proc. Int. Conf. Sound and Vibration (ICSV15), Daejeon, South Korea.
- Ku, E.M., Elliott, S.J. and Lineton, B. (2008). 'Modelling threshold fine structure and spontaneous otoacoustic emissions in the cochlea,' Proc. Brit. Appl. Maths Colloquium, Manchester, U.K.
- Ku, E.M., Elliott, S.J. and Lineton, B. (2008). 'Does the human cochlea work like a laser?' FESM Postgraduate Research Showcase, Southampton, U.K.
- Ku, E.M., Elliott, S.J. and Lineton, B. (2007). 'Modelling random and noise-induced changes in the parameters along the length of the cochlea and the effect on hearing sensitivity,' BSA Short Papers Meeting on Experimental Studies of Hearing and Deafness, London, U.K.
- Elliott, S.J., Ku, E. and Lineton, B. (2007). 'Some effects of spatial randomness along the length of the cochlear on its performance,' J. Acoust. Soc. Am. (153rd Meeting Acoust. Soc. Am.), 121, 3192.
- Ku, E. and Elliott, S.J. (2007). 'Comparing time domain simulations of different nonlinear models of cochlear micromechanics,' ARO Midwinter Research Meeting, Denver, USA.
- Elliott, S.J., Ku, E. and Lineton, B. (2007). 'The stability of a cochlear model assessed using a state space formulation,' ARO Midwinter Research Meeting, Denver, USA.
- Elliott, S.J. Ku, E. (2006). 'Feedback control of vibration in the inner ear,' Proc. International Symposium on Active Control of Sound and Vibration, Adelaide, Australia.

Chapter 2

the Neely and Kim (1986) Model

In 1986, S. T. Neely and D. O. Kim published a mathematical model of a cat cochlea which included a cochlear amplifier. As discussed in the introduction, this was among the first mature attempts in the field to explain the sharply tuned cochlear response by adding active feedback elements. Their formulation is based upon physical principles, anatomical characteristics, and observed responses of the cochlea. This physical interpretation is required in order to directly compare simulated responses with measurements. Through the course of this 3-year investigation, Neely and Kim's (1986) framework has been heavily studied, driven and tweaked. In order to account for the physical differences between the cat cochlea and a human cochlea, the lumped-element parameters describing the variation of the cochlea's mechanical properties as a function of position have been re-tuned.

This chapter is an in-depth examination of the Neely and Kim (1986) model, beginning with an overview of the entire system and its inherent assumptions. This is followed by a review of its passive formulation. The mechanism for applying an active feedback loop and its effect upon the passive system is then considered. Finally, the local motions of individual segments are coupled together by the cochlear fluid to give a simulation of the model's global behaviour.

2.1 Model Overview

The Neely and Kim (1986) model is an idealisation of the cochlea. The motions of the BM and TM are assumed to be linear, and the physical characteristics of the CP are lumped

into local masses, stiffnesses and dampers as a function of longitudinal position. Though it was originally formulated in the frequency domain, its mechanical basis lends itself well to simulation in the time domain (see Chapter 3). This representation of the cochlea can be discussed in terms of its *micromechanics* and its *macromechanics*. The term 'micromechanics' refers to the dynamic behaviour of a radial slice of the cochlea at the microscopic level (refer to Figure 1.4 for a review of the physiology). In contrast, the term 'macromechanics' deals with the coupling between the micromechanical motion of the system at various points along the cochlea, thus giving rise to a solution for the global response of the cochlea.

The micromechanical model of the cochlea is comprised of two masses, three springs and three dampers. This is illustrated in Figure 2.1 below. The model can be loosely interpreted as representing the anatomical features of a radial cross-section of the cochlea. For instance, the BM and the TM of the organ of Corti are modelled as masses (m_1 and m_2, respectively) which are coupled via the stiffness of the OHC stereocilia (k_3).

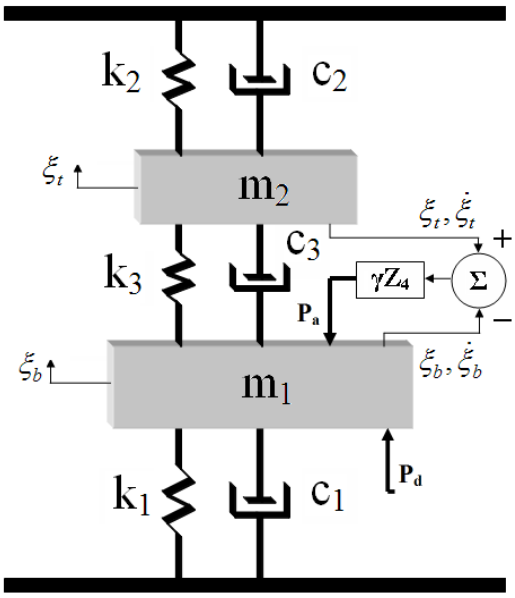


Figure 2.1: Neely & Kim's (1986) micromechanical model of the cochlea.

Note that this micromechanical model is not structurally coupled to adjacent slices of the cochlea; this feature is sometimes referred to as 'longitudinal coupling' in the literature (Robles and Ruggero, 2001).

The input to the micromechanical system comes in the form of a pressure applied upon the BM, p_d. This corresponds to a difference of pressures between the fluid-filled cavities of the cochlea above and below the CP. The resulting BM motion is detected by the CA, which is schematised as a feedback loop between the two masses. The CA is believed to activate when the ion-channel gating mechanisms of the OHC stereocilia are displaced due to the shearing motion of the TM relative to the BM. This in turn stimulates either hair bundle motility— the force due to the hair cell bundle, somatic motility— the force due to the lengthwise contractions and expansions of the OHCs, or both (Holley, 1996). The fine details of the electromechanical transduction process in the OC and CA are intentionally hidden in this model. A more detailed investigation of how the CA is modelled is given in section 2.3. One criticism of the Neely and Kim (1986) model has been that the active pressure source acting on the BM reacts against nothing (Hubbard and Mountain, 1996). Neely and Kim (1986) state that the active force 'pushes against the surrounding fluid;' however, this is not a very satisfactory explanation. As such, this represents a fundamental weakness of this model.

The micromechanical slices of the cochlea are assumed to be structurally independent, but the motions of nearby elements are coupled to one another via the cochlear fluid. A number of assumptions are inherent to this representation of the cochlear macromechanics. First, the cochlear fluid is defined as incompressible and inviscid (lossless). The incompressibility of the fluid disallows the existence of compression waves, waves within the cochlear fluid which travel at a high velocity. These compression waves are referred to as 'fast' waves in the literature (as opposed to the relatively 'slow' waves of local pressure difference which propagate along the CP). The importance of 'fast' waves in the cochlea is still being debated in the cochlear modelling community, though many of the salient features of the cochlea can be explained given only 'slow' TWs (Robles and Ruggero, 2001). In this work, all TWs are assumed to be 'slow' waves unless otherwise stated.

Another simplifying assumption regards the geometry of the cochlea. The human cochlea is curled into a spiral that typically exhibits two and a half turns (Pickles, 2003). In the Neely and Kim (1986) model, the cochlea is uncurled and modelled as rectangular box of length 35 mm, width 1 mm, and height 1 mm. This is shown below in Figure 2.2.

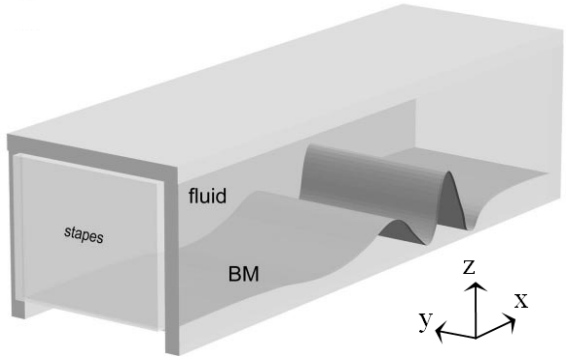


Figure 2.2: Schematic representation of Neely and Kim's 1-D macromechanical model of the cochlea: illustration of the exaggerated motion of cochlear segments due to sinusoidal excitation at the stapes. Reprinted from Hearing Research, 145, Kolston, P.J., 'The importance of phase data and model dimensionality to cochlear mechanics,' 25-36, Copyright (2000), with permission from Elsevier.

The stapes footplate, the component of the middle ear that is responsible for transmitting sound waves into the cochlea, is located at the oval window in the shaded face of Figure 2.2. The scala vestibuli and scala tympani, the fluid-filled tunnels above and below the basilar membrane, are simply modelled as fluids that rest above the CP. Located at the apex is the helicotrema, a passage that connects the two scalae. This allows for fluid flow from one chamber to the other.

Although the model has an assumed width, length and height, it is possible to reduce the mathematical analysis to a single dimension given the following assumption. As a tonally-generated TW propagates along the CP, its wavelength decreases monotonically until it approaches its frequency's resonant location (beyond this location, it is quickly extinguished). During most of the progress of the TW, its wavelength is long compared to the height- and width- dimensions of the scalae, hence the term, 'long-wave assumption.' The box model can thus be considered '1-D' over this range by excluding the vertical (z-axis) or transverse (y-axis) motion of the fluid from the mathematical analysis (de Boer, 1996). However, near and beyond the peak of the TW, this condition is violated; the implications of this are discussed in Chapter 6.

2.2 Passive Micromechanics

The micromechanical model of the cochlea consists of a two-degree-of-freedom system (see Figure 2.1) which represents a discrete radial slice of the cochlea. The top and bottom fixed points refer to the rigid boundary of the cochlear walls. The relative motion of m_1

and m_2 corresponds to the shear displacement between the BM and the TM. The springs and dampers in the model represent the biological stiffnesses and compliances of the cochlea and the fluid surrounding the OC. The values of these elements are discussed in the next section.

2.2.1 Variation of model parameters with longitudinal position

The micromechanical framework of Neely & Kim's model of the cochlea gives rise to solutions for the motion of the BM and TM given a local input pressure. The parameters for the model stiffnesses and compliances that vary with position along the BM were originally chosen to reflect the tonotopy of a cat cochlea. As discussed previously, these values have been re-tuned to match certain characteristics of the human cochlea. It should be noted that there have been three sets of parameters used through the course of this investigation: 1) Neely and Kim's (1986) parameters for the cat cochlea; 2) a revised set of parameters for the human cochlea, published in Ku *et al*. (2008); and 3) a further refined set of human parameters. The last set of values is presented here, and all of the results in the thesis are derived from this model. The data published in Elliott *et al.* (2007) and Ku *et al.* (2008), included in Appendix C, are derived from sets 1) and 2), respectively.

The refined values are compared to the original values in Table 2.1. The numbered-subscript quantities correspond to micromechanical elements, while *m*-subscript quantities refer to the middle ear boundary and x is the longitudinal distance along the cochlea. γ is the micromechanical feedback gain, while g, b, L and H are the BM to IHC lever gain, the ratio of the maximum to the average vertical displacement of the BM across one radial slice, the length of the cochlea, and the height of the fluid channel. Finally, c_h is the damping at the helicotrema, A_s is the area of the stapes footplate, ρ is the density of the cochlear fluid, and N is the number of elements in the model. The variation of the values of these revised parameters is plotted against longitudinal position along the cochlea in Figure 2.3.

Quantity	Original Formula (SI)	Revised Formula (SI)	Units
$k_1(x)$	$1.1\times10^{10}e^{-400x}$	$1.65\times10^{9}e^{-279(x+0.00373)}$	$N\times m^{-3}$
$c_1(x)$	$200+15000\ e^{-200x}$	$9+9990e^{-153(x+0.00373)}$	$N\times s\times m^{-3}$
$m_1(x)$	3×10^{-2}	4.5×10^{-3}	$kg\times m^{-2}$
$k_2(x)$	$7\times10^{7}e^{-440x}$	$1.05\times10^{7}e^{-307(x+0.00373)}$	$N\times m^{-3}$
$c_2(x)$	$100\ e^{-220x}$	$30e^{-171(x+0.00373)}$	$N\times s\times m^{-3}$
$m_2(x)$	5×10^{-3}	$7.20\times10^{-4}+2.87\times10^{-2}\boldsymbol{x}$	$kg\times m^{-2}$
$k_3(x)$	$1\times10^{8}e^{-400x}$	$1.5\times10^{7}e^{-279(x+0.00373)}$	$N\times m^{-3}$
$c_3(x)$	$100e^{-80x}$	$6.6e^{-59.3(x+0.00373)}$	$N\times s\times m^{-3}$
$k_4(x)$	$6.15\times10^{9}e^{-400x}$	$9.23\times10^{8}e^{-279(x+0.00373)}$	$N\times m^{-3}$
$c_4(x)$	$100e^{-80x}$	$3300e^{-144(x+0.00373)}$	$N\times s\times m^{-3}$
γ	1	1	--
g	1	1	--
b	0.4	0.4	--
L	0.025	0.035	m
H	0.001	0.001	m
k_m	2.1×10^{6}	2.63×10^{8}	$N\times m^{-3}$
c_m	4000	2.8×10^{4}	$N\times s\times m^{-3}$
m_m	45×10^{-2}	2.96×10^{-2}	$kg\times m^{-2}$
c_h	0	350	$N\times s\times m^{-3}$
A_s	1×10^{-6}	3.2×10^{-6}	m^2
ρ	1000	1000	$kg\times m^{-3}$
N	251	500	--

Table 2.1: Model parameters for Neely and Kim's (1986) cat cochlea (second column), and the revised quantities for the human cochlea (third column).

Neely and Kim (1986) note that their values 'were selected to simulate the biomechanics of a cat cochlea with consideration given to the physical structure, frequency-to-place map, and frequency tuning curves typical for a cat.' Similarly, the revised values were chosen such that the global response of the model exhibits observed and inferred TW properties of the human cochlea such as tonotopy, enhancement characteristics, and TW wavelength as a function of position. These quantities of the coupled cochlea are presented in section 2.4.

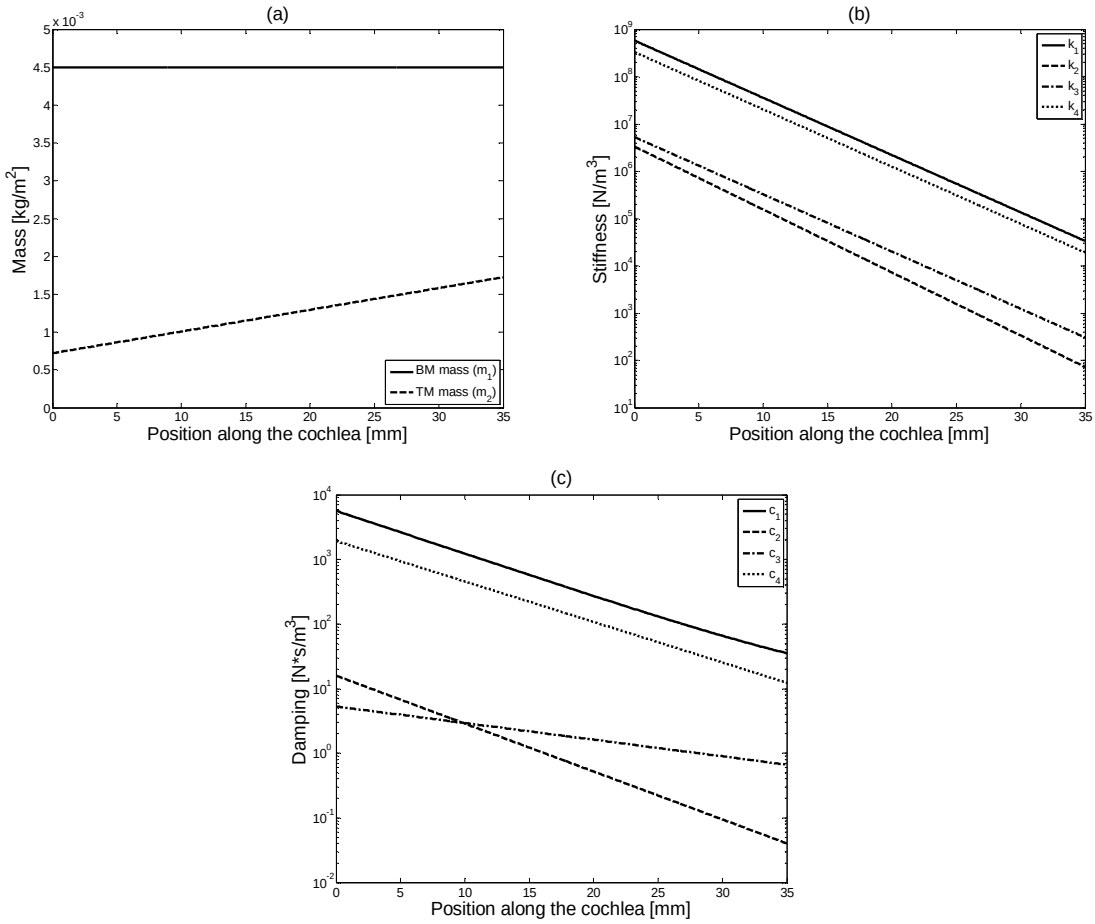


Figure 2.3.a-c: Variation of model parameters with position.

As with the original set of parameters, the stiffness and damping parameters all decrease exponentially as a function of position from the base of the cochlea. However, the mass of the TM now increases linearly from 0.72 g/m^2 to 1.7 g/m^2, whereas the mass of the BM is still held constant. This modification was necessary in order to increase the amount of amplification toward the apex relative to the amplification provided by a constant TM mass. In addition, the overall values of many non-damping parameters were reduced by almost a factor of 7 relative to the original Neely and Kim (1986) values in order to shorten the wavelength of the TW. The values of the damping terms were reduced to a lesser degree to maintain a plausible amount of active gain.

2.2.2 Micromechanical frequency response functions

Each micromechanical model represents the averaged motion of that slice of the cochlea in the radial direction. Neely and Kim define *b* as the ratio of the average displacement

across the width of the CP, ξ_p, to the maximum displacement over the width of the BM, ξ_b, so that

$$\xi_p(x) = b\xi_b(x). \tag{2.1}$$

As the input to the system is the local pressure difference, p_d, it is instructive to calculate the BM velocity for a given pressure; the ratio of output velocity to input pressure is defined as mobility or admittance. The transfer function of the BM admittance $\frac{\dot{\xi}_b(x)}{p_d(x)}$, as given by Neely and Kim and derived in Appendix B, is

$$\frac{\dot{\xi}_b(x)}{p_d(x)} = \frac{1}{g\left[Z_1(x) + Z_2(x)\left(\frac{Z_3(x) - \gamma Z_4(x)}{Z_2(x) + Z_3(x)}\right)\right]}, \tag{2.2}$$

where

$$\begin{aligned} Z_1(x) &= k_1(x)/s + c_1(x) + sm_1(x) \\ Z_2(x) &= k_2(x)/s + c_2(x) + sm_2(x) \\ Z_3(x) &= k_3(x)/s + c_3(x) \\ Z_4(x) &= k_4(x)/s + c_4(x) \end{aligned}, \tag{2.3}$$

g is the BM to IHC lever gain and here, $s \equiv j\omega$. Z_1 represents the mechanical impedance of the organ of Corti; Z_2 represents the mechanical impedance of the tectorial membrane; Z_3 represents the coupling between the OC and the TM; and Z_4 represents the impedance associated with the active pressure source. The term γ denotes feedback gain, where $\gamma = 0$ generates a passive response, and $\gamma = 1$ generates a baseline active response. The term 'baseline' is used in this thesis to describe an active, passive, linear or nonlinear model without perturbations.

The FRF of the TM can be expressed as a function of the BM FRF. The solution for the TM mobility, or admittance, is:

$$\frac{\dot{\xi}_t(x)}{p_d(x)} = \frac{1 - \left[\dot{\xi}_b(x)/p_d(x)\right]\left[gZ_1(x)\right]}{\left[Z_2(x) + \gamma Z_4(x)\right]}, \tag{2.4}$$

where $\dot{\xi}_t$ is the TM velocity.

2.2.3 Passive admittance

The two degree-of-freedom system that Neely & Kim use to represent the micromechanical behaviour of the cochlea has two modes of vibration. In order to simplify the analysis, the values of the dampers in the system are first reduced by a factor of 1000. For the lower-frequency mode, the motions of the BM and the TM are in-phase. For the higher-frequency mode, the motions of the masses are out-of-phase. This can readily be observed by examining the near-undamped magnitudes and phases of the BM and TM admittance for a micromechanical element at a single position along the cochlea, as given in Figure 2.4.a, c.

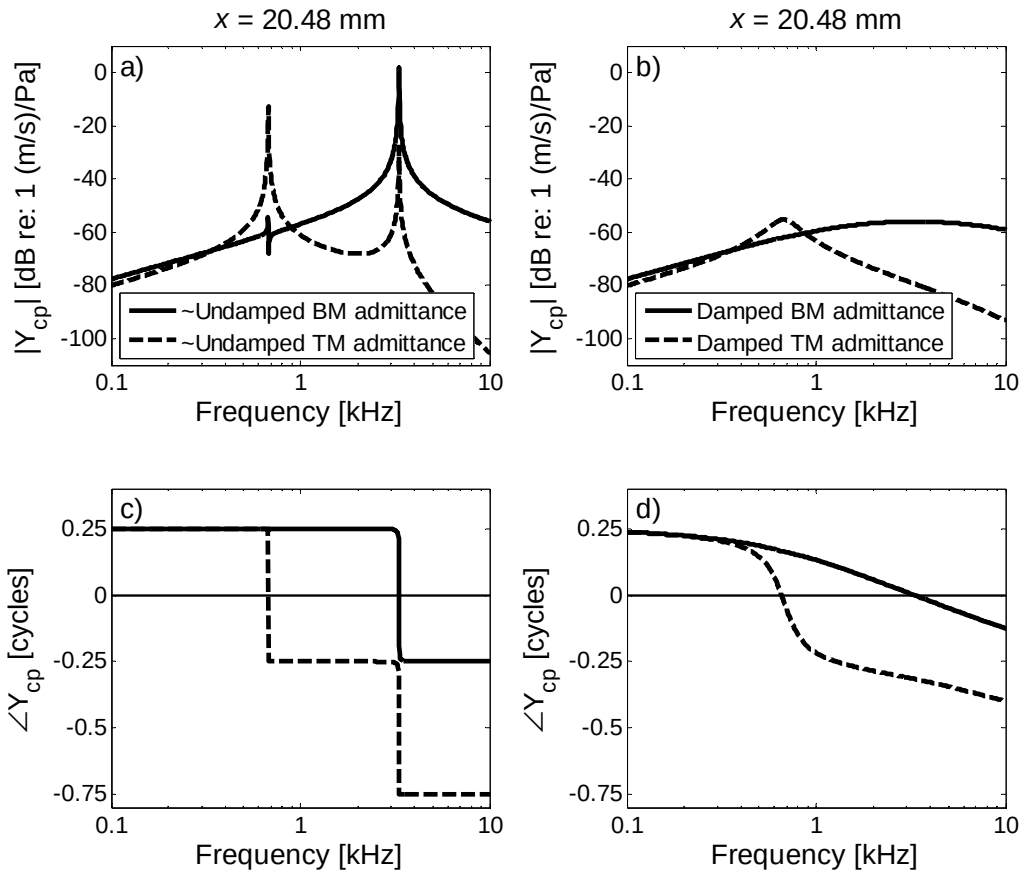


Figure 2.4.a-d: Magnitudes and phases of the damped and near-undamped system admittance at x = 20.48 mm.

With the damping present in the system (b and d panels), the sharpness of the peaks is significantly reduced. This is particularly visible at the higher resonant frequency. The

motion of the two masses remains close to 180 degrees (or half a cycle) out-of-phase at frequencies higher than that of the first resonant peak.

2.2.4 Undamped natural frequencies

In order to better understand the model's micromechanical response as a function of position along the BM, it is useful to investigate the undamped natural frequencies of the system. The resonant modes of this two degree-of-freedom system are derived in Appendix B, thus showing that:

$$\omega^2 = \left(\frac{k_1+k_3}{2m_1}+\frac{k_2+k_3}{2m_2}\right) \pm \sqrt{\left(\frac{k_1+k_3}{2m_1}+\frac{k_2+k_3}{2m_2}\right)^2 - \frac{\left(k_1k_2+k_1k_3+k_2k_3\right)}{m_1m_2}}\,. \qquad (2.5)$$

This equation returns two positive solutions for ω^2 which correspond to the solutions for ω_1 and ω_2.

A simplified solution which very closely approximates the full analytical solution can be found by simplifying the motion of the masses at each mode, as shown in Figure 2.5.

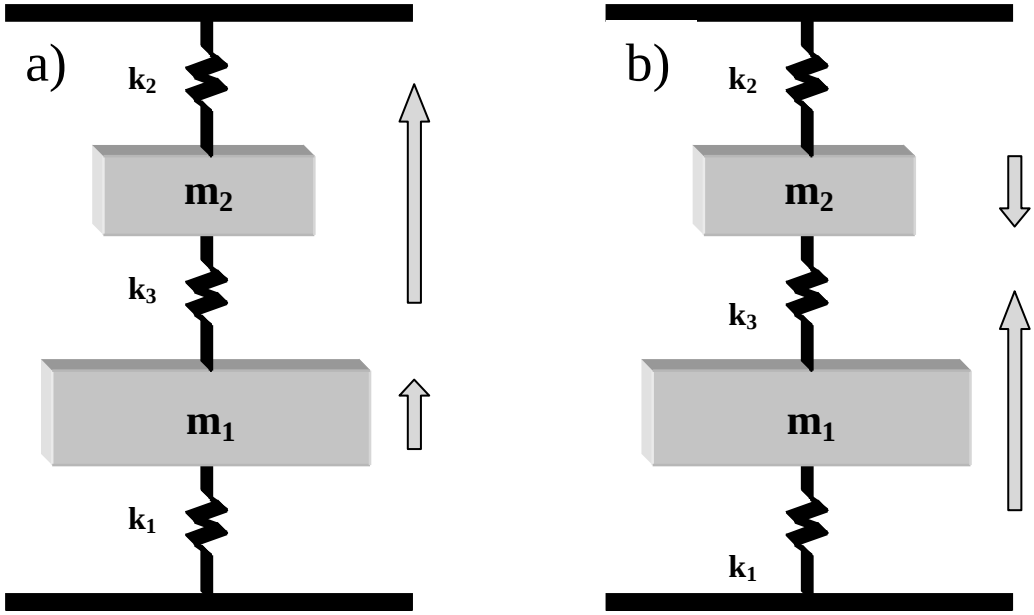


Figure 2.5.a-b: Illustration of the relative magnitudes and directions of motion of the BM (m_1) and the TM (m_2) at the first (left panel) and second (right panel) modes of oscillation.

In the first mode of oscillation, the BM moves much less than the TM. As such, the BM can be considered stationary for the purposes of this exercise. Furthermore, the value of

the springs k_2 and k_3 in this model are very similar. Subsequently, the frequency of the first mode can be simplified to

$$\omega_1 \approx \sqrt{\frac{k_2 + k_3}{m_2}}. \tag{2.6}$$

In the case of the second mode, the motion is dominated by the BM. In addition, the stiffness of k_1 is much greater than k_2 or k_3. This suggests that the frequency of the second mode can be approximated by

$$\omega_2 \approx \sqrt{\frac{k_1}{m_1}}. \tag{2.7}$$

The variation of ω_1 and ω_2 is plotted against position below in Figure 2.6. The resonant frequencies decrease exponentially as a function of position along the BM. The approximate results of equations (2.6) and (2.7) are indistinguishable from the exact results of equation (2.5) when plotted on the scale shown in Figure 2.6.

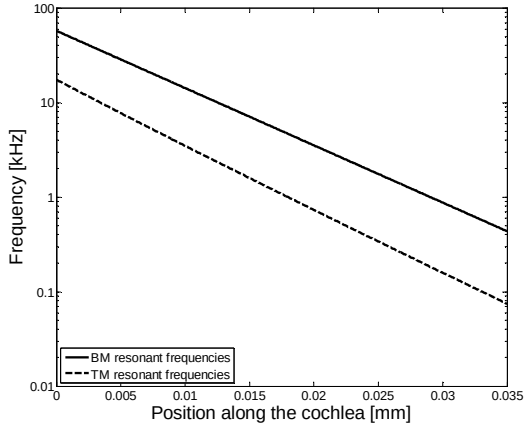


Figure 2.6: Undamped, micromechanical BM and TM resonant modes as a function of position along the BM.

2.3 Active Micromechanics

By definition, active amplification involves reducing the effective or observed losses in a system. An electrical circuit example of amplification might consist of resistors and an operational amplifier. The op-amp increases the current or voltage at the output, relative to what could be supplied by the input source alone. In a mechanical analogue of this system, the electrical resistors would be represented as dampers. Thus, in the cochlea, it is believed that the CA overcomes the damping present in the fluid by mechanically amplifying the motion of the TW as it propagates along the BM. This 'undamping' allows

the IHCs (at the output stage of the mechanical system) to detect a stronger signal (IHC stereocilia motion due to fluid drag in the sub-tectorial space) (Robles and Ruggero, 2001).

The analysis of a tonally driven system is often complex, literally. For instance, a circuit may include energy storage elements such as inductors and capacitors, or in the mechanical analogue, masses and springs. The mechanical admittance of such a system will include real (dissipative) and imaginary (energy storing and releasing) components. In a completely passive system, the real part of its admittance will be positive across frequency so that it only absorbs energy. However, in an active system, the real part of its admittance will be reduced in a given frequency range and may even be negative, thus indicating that the system can supply energy. When point measurements of the admittance of the mammalian cochlear partition are made, it is found that there are indeed frequency bands with a negative-real parts suggesting an active process; this gives rise to the term 'negative damping' in the literature (de Boer, 1996).

2.3.1 Active admittance

As discussed previously, the Neely and Kim micromechanical model is driven by an active mechanism which represents the contribution of the outer hair cells. In their 1986 paper, Neely and Kim state that the active impedance, Z_4, is 'included to provide a frequency-dependent phase shift' between the active pressure source and the relative motion of the BM and TM. This is given by

$$p_a(x) = -\gamma Z_4(x)\dot{\xi}_c(x), \tag{2.8}$$

where Z_4 consists of both real and imaginary components, p_a, the pressure generated by the outer hair cells, and $\dot{\xi}_c$ is defined as the difference between the TM and BM velocities:

$$\dot{\xi}_c = g\dot{\xi}_b - \dot{\xi}_t(x). \tag{2.9}$$

The real component of Z_4 primarily serves to reduce the damping of the CP, while the imaginary component provides a shift in frequency relative to the passive resonances. The feedback function takes as its input the difference in displacements and velocities of the BM and TM. The output is p_a, the pressure generated by the outer hair cells; the magnitude of this output is scaled by a unitless quantity, γ, which is nominally set to unity.

The active admittance of Neely and Kim's BM is again

$$\frac{\dot{\xi}_b(x,\omega)}{p_d(x,\omega)} = \left\{ g\left[Z_1(x) + Z_2(x)\left(\frac{Z_3(x) - \gamma Z_4(x)}{Z_2(x) + Z_3(x)} \right) \right] \right\}^{-1}. \tag{2.2}$$

Both the active ($\gamma = 1$) and the passive ($\gamma = 0$) admittance are plot as a function of position (for a single frequency) in Figure 2.7. Similarly, these quantities are plot as a function of frequency (at a single position) in Figure 2.8. As expected, the real part of the admittance (b panels) is reduced in a particular region in frequency and position along the cochlea. However, at locations apical to the negative-damping region, the active admittance increases the damping for a 1 kHz stimulus. In the coupled cochlea, this would serve to sharpen the response of the TW spatially by actively attenuating BM motion beyond the resonant location.

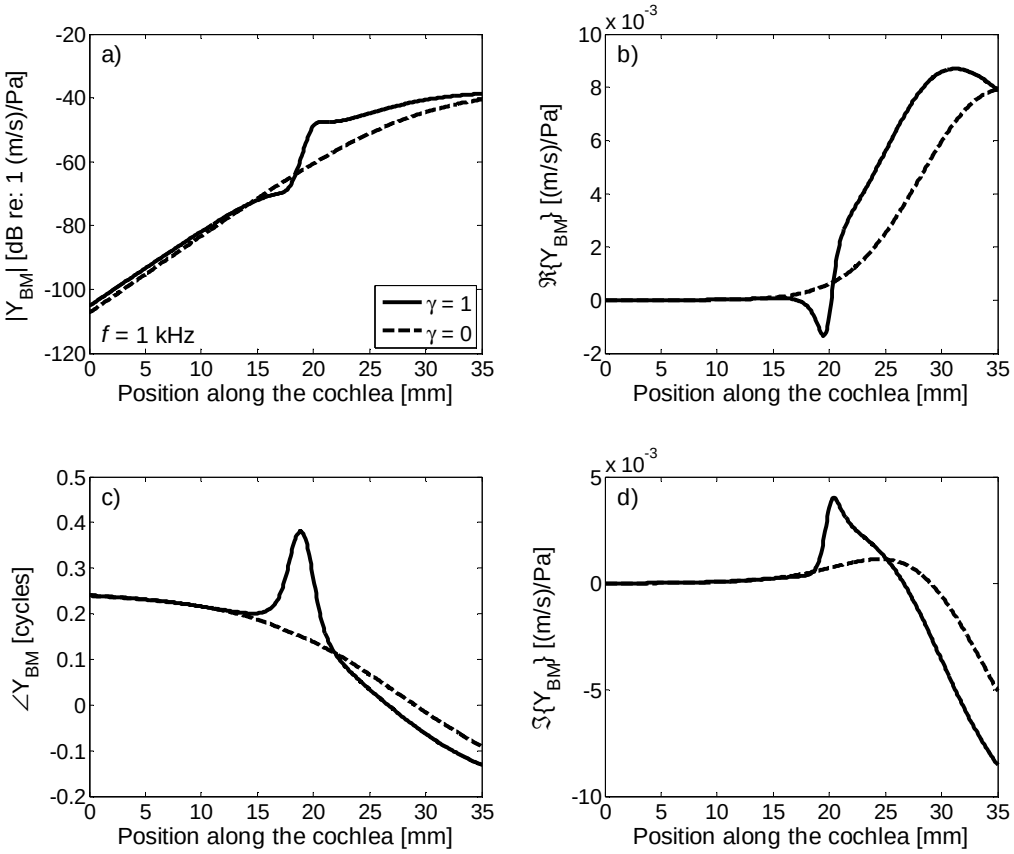


Figure 2.7.a-d: Active BM admittance magnitude (a) and phase (b) as a function of position along the cochlea given a 1 kHz excitation. The real (b) and imaginary (d) parts of the BM admittance are also presented. Active responses ($\gamma = 1$) are shown with a solid line, whereas passive responses ($\gamma = 0$) are shown with a dashed line.

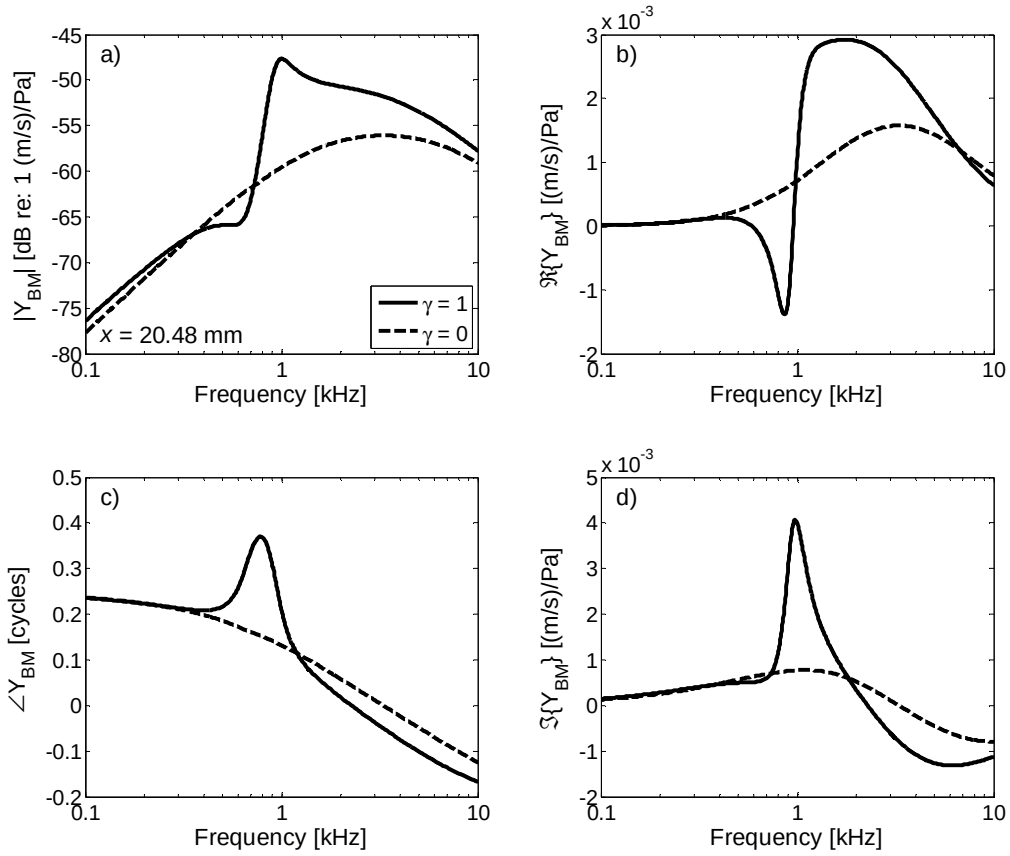


Figure 2.8.a-d: Active BM admittance magnitude (a) and phase (b) for the isolated micromechanical model as a function of frequency at a location 20.48 mm from the base. The real (b) and imaginary (d) parts of the BM admittance are also presented. Active responses ($\gamma = 1$) are shown with a solid line, whereas passive responses ($\gamma = 0$) are shown with a dashed line.

2.3.2 Micromechanical stability analysis

If a linear system includes active feedback, there is the possibility that it may become unstable such that its response grows without bound. Analyzing the stability of the micromechanical model can provide insight into the behaviour of the coupled cochlea. It has been theorised that the spontaneous emission of sound from the cochlea is indicative of unstable oscillators in the cochlea (e.g. Duke and Jülicher, 2003). It should be noted, however, that the stability of the individual micromechanical elements gives no guarantee of the stability of the coupled cochlea, as explained in more detail in the next chapter. Furthermore, though it is possible to determine the frequency response of an unstable system, such a result does not have physical significance and may lead to the

misinterpretation of calculations. Thus, it is necessary to determine if isolated micromechanical elements are stable or not.

According to classical control theory, a system that includes active feedback can be classified in terms of its plant (*G*), and its feedback mechanism (*H*). A Nyquist Diagram is a plot of the product of these quantities on real and imaginary axes, which varies with frequency. A general and isolated feedback loop is shown in Figure 2.9.

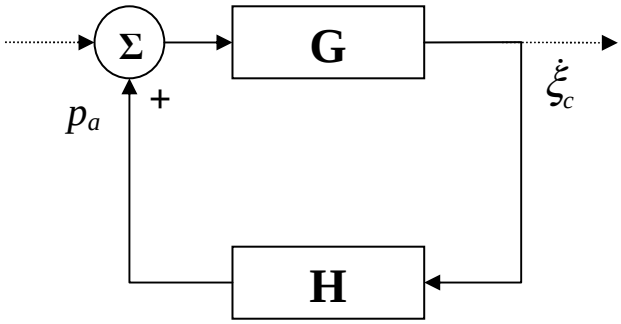


Figure 2.9: General diagram of a feedback loop containing a micromechanical controller. The dotted lines represent the system input and output paths.

For the Neely and Kim system, the plant *G* and feedback mechanism *H* are derived in Appendix B and are given by:

$$G = \frac{\dot{\xi}_c(x)}{p_a(x)} = \left(\frac{Z_2}{Z_1 Z_2 + Z_1 Z_3 + Z_2 Z_3} \right), \tag{2.10}$$

and

$$H = -\gamma Z_4(x). \tag{2.11}$$

The quantity *GH* is referred to as the open-loop transfer function:

$$GH = -\gamma Z_4 \left(\frac{Z_2}{Z_1 Z_2 + Z_1 Z_3 + Z_2 Z_3} \right). \tag{2.12}$$

The open-loop function can provide a means of determining the closed-loop stability of an active-feedback system; this is accomplished using the Nyquist Stability Theory (Phillips and Harbor, 2000). A Nyquist Diagram is a plot that shows the real and imaginary components of *GH* as they vary with frequency; this is shown for the micromechanical system in Figure 2.10.

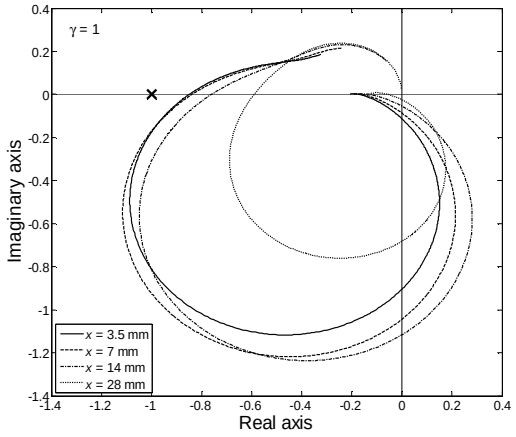


Figure 2.10: Plot of the open-loop transfer function, *GH*, for the micromechanical oscillators at several positions along the cochlea. An 'x' marks the -1+*j*0 location. Only positive frequencies from 20 Hz to 50 kHz are shown for clarity.

It is clear from Figure 2.10 that some locations along the cochlea operate quite close to instability at a gain of $\gamma = 1$. In general, for a given gain, the system becomes more stable as one examines positions further towards the apex of the cochlea. As a system becomes more stable, its response is less strongly amplified.

Given the open-feedback loop of a system, it is sufficient to say that the system is unstable if $GH(j\omega)$ crosses the real axis at or below the -1+j0 point on the Nyquist Diagram. The maximum stable gain of a system is thus

$$\gamma_{\max} = \frac{-1}{GH\,|_{\gamma=1,\mathrm{Im}\{GH\}=0}}. \tag{2.13}$$

In classical control theory, the closed-loop enhancement provided by a feedback controller is given by:

$$Enhancement(dB) = 20\log_{10}\left(\frac{1}{1+GH\,|_{\gamma=1,\mathrm{Im}\{GH\}=0}}\right). \tag{2.14}$$

Equivalently, when equation (2.13) is combined with (2.14),

$$Enhancement(dB) = 20\log_{10}\left(\frac{\gamma_{\max}}{\gamma_{\max}-1}\right). \tag{2.15}$$

Figure 2.11 displays the minimum gain values before the isolated micromechanical system is driven into instability as a function of position along the cochlea, in addition to the predicted enhancement provided by the active mechanism with $\gamma = 1$.

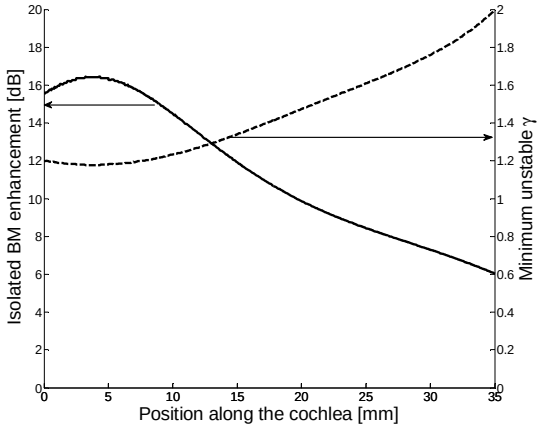


Figure 2.11: Isolated enhancement given $\gamma = 1$ (solid line, left axis) and minimum values of gain, γ, that lead to instability for various positions along the cochlea (dashed line, right axis).

The predicted values of gain that lead to instability can be tested by examining the frequency response of a system with different gains. At a location of 20.48 mm, the stability analysis predicts that the model will become unstable at $\gamma = 1.485$. Figure 2.12 shows the admittance at this location given several values of γ above and below this value. While the change in the magnitude of the calculated admittance is almost imperceptible when the gain is increased from $\gamma = 1.48$ to $\gamma = 1.49$, the sign of the phase flips suddenly above the resonant frequency; this indicates that the system is unstable.

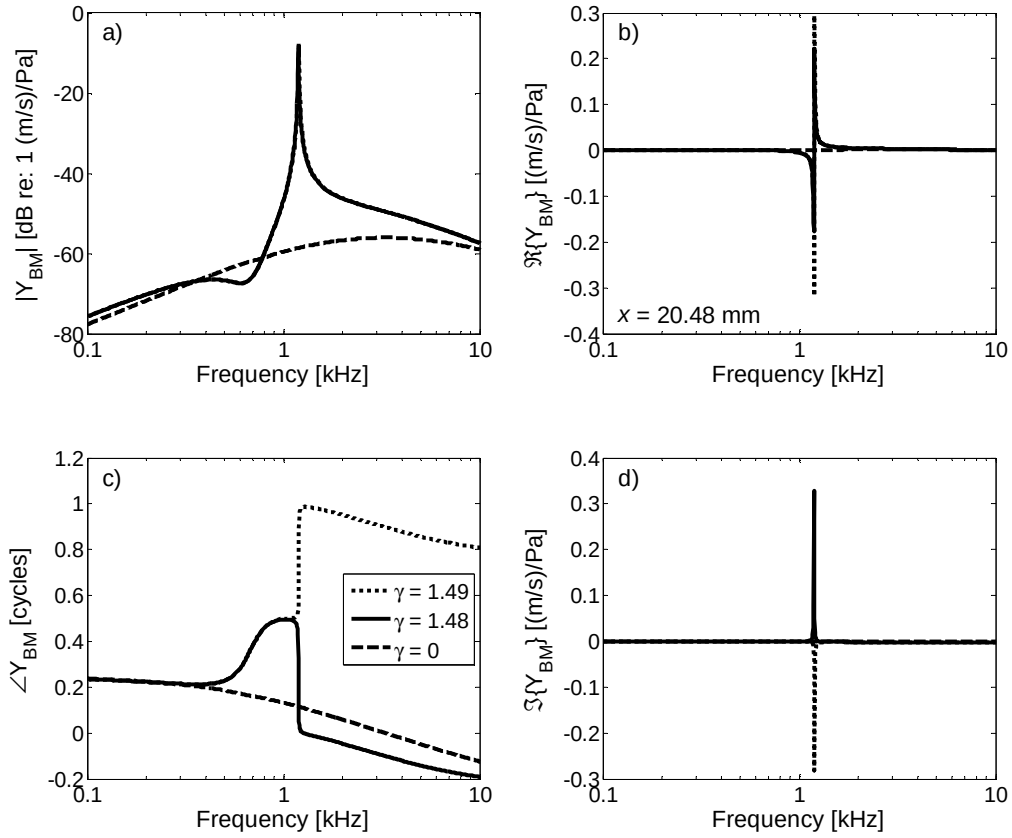


Figure 2.12: Calculated admittance of an isolated micromechanical element as a function of frequency at x = 20.48 mm for 3 values of γ. The magnitude (a), phase (c), real (b) and imaginary (d) parts of the BM admittance are all presented. Active and stable responses (γ = 1.48) are shown with a solid line, active and unstable responses (γ = 1.49) are shown with a dotted line, and passive responses (γ = 0) are shown with a dashed line.

2.4 Cochlear Macromechanics

Up until now, the motion of the cochlea has only been considered at the microscopic level. The conventional macromechanical model of the cochlea can be visualised as a series of independent oscillators that are coupled through the fluid. This idea is illustrated in Figure 2.13.

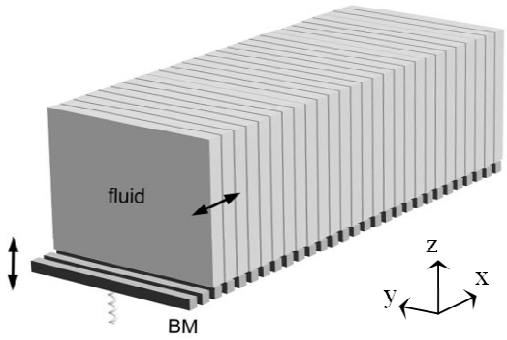


Figure 2.13: Schematic of a 1D box model of cochlear macromechanics: detail view of the conceptual construction of the macromechanics. The BM spans the entire width of the cochlear duct, and segments are coupled only by the fluid. Reprinted from Hearing Research, 145, Kolston, P.J., 'The importance of phase data and model dimensionality to cochlear mechanics,' 25-36, Copyright (2000), with permission from Elsevier.

The absence of direct structural coupling between the micromechanical elements is a characteristic feature of 'classical' 1-D models of the cochlea (de Boer, 1996). A number of scientists have made arguments for various forms of 'longitudinal coupling' through the mechanical structures of the cochlea (e.g. Fukazawa, 2002), though many of the salient features of the cochlea can be generated in a 'classical' model where the micromechanics are locally reacting. Greater attention is given to this point in the discussion, Chapter 6.

The longwave approximation assumes that the energy of a pressure wave only propagates in the longitudinal (*x*-) direction along the cochlea, and not in radial (*y*-) or vertical (*z*-) directions. This permits the reduction of the order of the model to a single dimension. As a result, a one-dimensional wave equation can be written in terms of the differential pressure across the cochlear partition, p_d:

$$\frac{\partial^2 p_d(x,\omega)}{\partial x^2} + \kappa_{TW}{}^2(x,\omega)\, p_d(x,\omega) = 0\,, \tag{2.16}$$

where κ_{TW} is the wavenumber of the travelling wave. The wavenumber is a quantity with units of inverse distance and is the spatial analogue of frequency. For instance, a TW with

a short wavelength would have a high wavenumber, just as a tone with a short period would have a high frequency. The wavenumber is related to the impedance of the cochlear partition by

$$\kappa_{\mathrm{TW}}{}^{2}(x,\omega) = \frac{-2j\omega\rho}{HZ_{cp}(x,\omega)}, \tag{2.17}$$

where H is the height of the fluid chamber, and ρ is the density of the cochlear fluid. A full derivation of the wave equation in a cochlear transmission line can be found, for example, in de Boer (1991). A boundary condition is set at the base of the cochlea:

$$\left.\frac{\partial p}{\partial x}\right|_{x=0} = -2j\omega\rho u_{st}, \tag{2.18}$$

where the stapes velocity (u_{st}) is specified. Similarly, the boundary condition at the helicotrema is given in this model by:

$$p_d\big|_{x=L} = 0. \tag{2.19}$$

It should be noted that in most 1-D formulations, as described above, the apical boundary condition is simply left as a pressure release. It can be modified to allow for a small degree of damping to be added at the helicotrema, thus reducing apical reflections at low frequencies; this is described in Appendix C.2.

2.4.1 Numerical implementation of frequency domain macromechanics

The macromechanical dynamics are modelled here following the methodology outlined by Neely and Kim (1986) and Neely (1981). The length of the cochlea (*L*) is divided into *N* elements, each of which has a length Δ, where Δ=*L*/(*N*-*1*). A finite difference method is applied to adapt the spatially continuous wave equation and boundary conditions to describe a discrete number of points. The wave equation (2.16) can be rewritten using the Taylor series expansion for the second derivative of p_d with respect to *x*:

$$\frac{p_d(n+1) - 2p_d(n) + p_d(n-1)}{\Delta^2} - \frac{2j\omega\rho}{HZ_{cp}(n)}p_d(n) = 0, \tag{2.20}$$

where *n* denotes which element of the 1D model is being evaluated. For instance, the cochlear elements are represented by *n* = 2, 3, …, *N*-1. The boundary condition at the base, (2.18), can also be written using another finite difference approximation:

$$\frac{p_d(2) - p_d(1)}{\Delta} = -2j\omega\rho u_{st}. \tag{2.21}$$

At the apex, (2.19) is rewritten as

$$p_d(N) = 0. \tag{2.22}$$

The wave equation and system boundary conditions can be expressed in matrix form:

$$(\mathbf{F} - \mathbf{M})\mathbf{p_d} = \mathbf{q}, \tag{2.23}$$

where bold font indicates the quantity is a matrix. **F** is the tri-diagonal $N \times N$ element fluid-coupling matrix,

$$\mathbf{F} = \frac{1}{\Delta^2}\begin{bmatrix} -\Delta & \Delta & & & 0 \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ 0 & & & & \Delta^2 \end{bmatrix}; \tag{2.24}$$

M is the diagonal $N \times N$ element mobility (admittance) matrix,

$$\mathbf{M} = \frac{2j\omega\rho}{H}\begin{bmatrix} Y_m & & & & 0 \\ & Y_{cp}(2) & & & \\ & & \ddots & & \\ & & & Y_{cp}(N-1) & \\ 0 & & & & Y_H \end{bmatrix}, \tag{2.25}$$

where $\mathbf{p_d}$ is an $N \times 1$ matrix of local pressure differences,

$$\mathbf{p_d} = \begin{bmatrix} p_d(1) \\ \vdots \\ p_d(n) \\ \vdots \\ p_d(N) \end{bmatrix}; \tag{2.26}$$

and **q** is the $N \times 1$ input matrix,

$$\mathbf{q} = \begin{bmatrix} -2j\omega\rho u_{st} \\ 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix}. \tag{2.27}$$

The **F** and **M** matrices can be combined to yield a single tri-diagonal matrix T, such that

$$\mathbf{T} = \mathbf{F} - \mathbf{M}, \tag{2.28}$$

which, combined with equation (2.23), yields

$$\mathbf{T}\mathbf{p_d} = \mathbf{q}. \tag{2.29}$$

The distribution of pressure differences is then obtained by inverting **T**:

$$\mathbf{p_d} = \mathbf{T}^{-1}\mathbf{q}. \tag{2.30}$$

Note that the matrix-inversion is evaluated in MATLAB by applying the left-division operator; this performs the inversion by Gaussian elimination (Neely and Kim, 1986).

The BM velocity at any position can thus be calculated by taking the quotient of the pressure difference at that location and the local CP impedance,

$$\dot{\xi}_b(x,\omega) = \frac{p_d(x,\omega)}{Z_{cp}(x,\omega)}. \tag{2.31}$$

The BM displacement is then found by integrating in the Laplace domain:

$$\xi_b(x,\omega) = \frac{\dot{\xi}_b(x,\omega)}{s}. \tag{2.32}$$

2.4.2 Response of the coupled cochlea

This section presents some typical responses of the coupled cochlea to tonal stimuli. The first plot, Figure 2.14, shows the pressure difference and BM velocity as a function of position for generated by a 1 kHz excitation in Neely and Kim's coupled model.

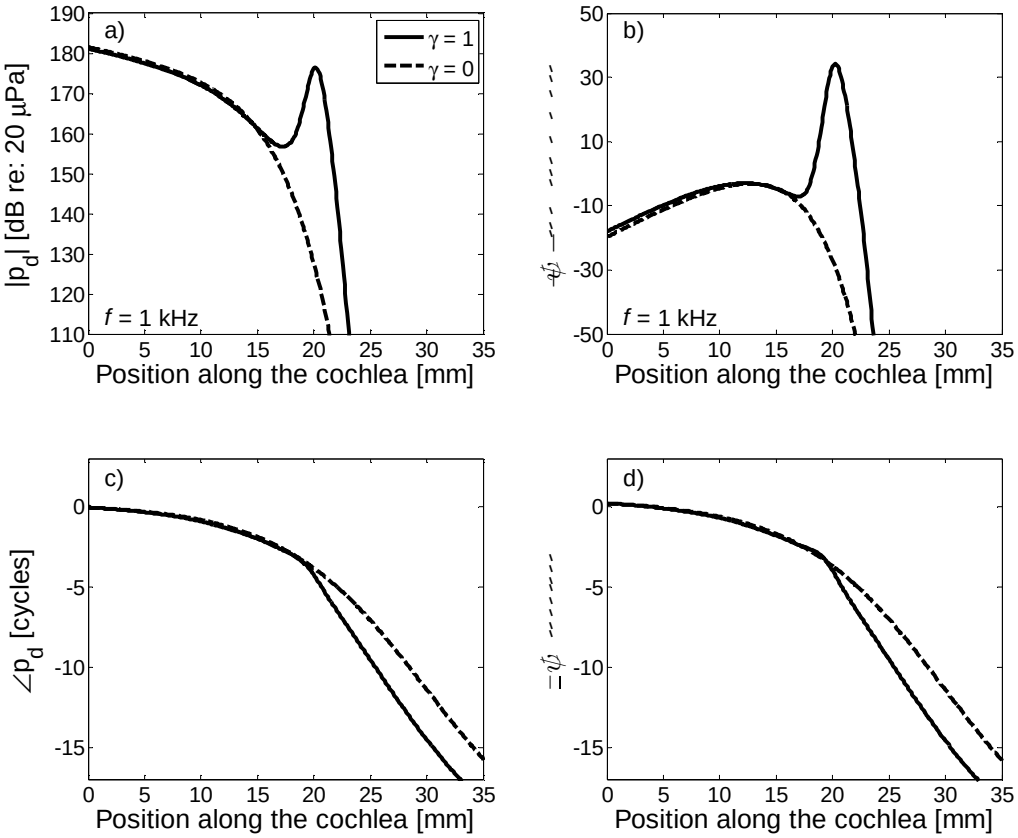


Figure 2.14.a-d: Pressure (a, c) and BM velocity (b, d) magnitude (a, b) and phase (c, d) given a 1kHz stimulus tone in the coupled cochlea given active (solid line, $\gamma = 1$) and passive (dashed, $\gamma = 0$) models.

It is clear from both the pressure and the velocity response that the sharpness of the tuning is increased when the cochlea is active. The phase of the response is approximately the same for both active and passive cases up until the peak of the response, where the active case lags behind the passive case by approximately half a cycle. This behaviour is also visible when the response is plotted as a function of frequency at a single location, as in Figure 2.15.

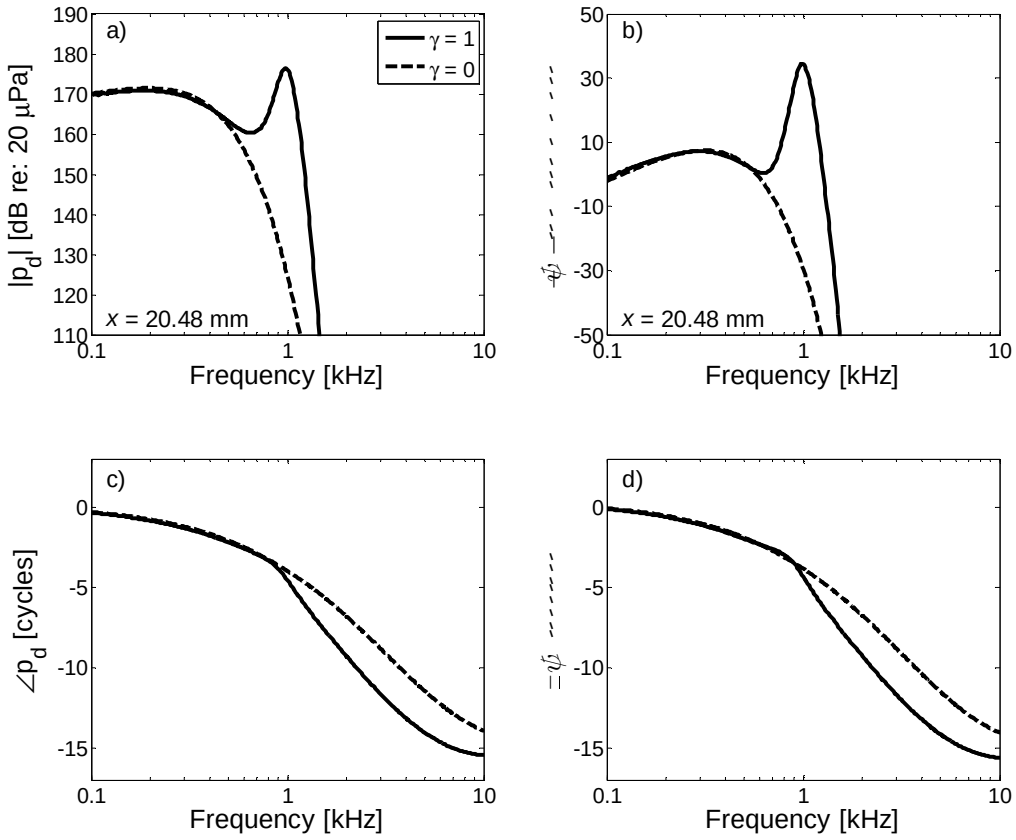


Figure 2.15.a-d: Pressure (a, c) and BM velocity (b, d) magnitude (a, b) and phase (c, d) as a function of frequency at the 1kHz characteristic place in the coupled cochlea given active (solid line, $\gamma = 1$) and passive (dashed, $\gamma = 0$) models.

The gain in the magnitude of the response at a given location along the mammalian cochlea is one of the quantities that is commonly measured *in vivo*. According to Robles and Ruggero (2001), the gain in BM motion provided by the CA of a chinchilla is approximately 45 dB at the base, and decreases to ~20 dB toward the apex. Figure 2.16 plots the enhancement as a function of position for the model used here. The enhancement

is defined as the quotient of the maximum active velocity and the maximum passive velocity at a given position along the cochlea, across frequency.

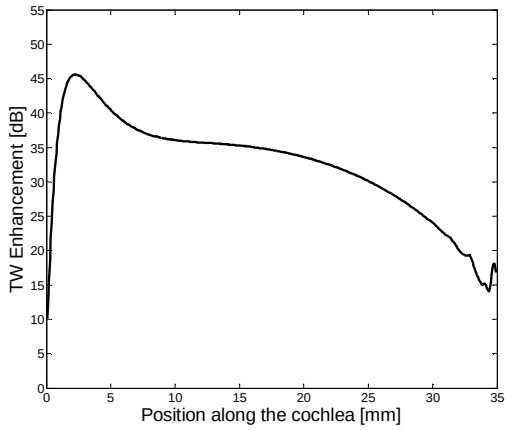


Figure 2.16: Enhancement of TW motion in the coupled cochlea.

Note the shape of Figure 2.16, in particular the sharp rise in enhancement near the base. This is due to the widening region of active undamping of the TW with decreasing frequency; at the very base, the TW is not strongly amplified because the spatial region of active OHC contribution is small. This increases to a nominal amount at approximately x = 3 mm, after which the other characteristics of the model dominate the enhancement curve. The shape of the enhancement curve is in qualitative agreement with published physiological data, where it is approximately 45 dB at the base and decreases to ~20 dB at the apex. The exact motion of the BM in the middle of the cochlea has not yet been measured due to experimental restrictions; until further advancements in measurement techniques are made, it will be difficult to validate the model in this region.

Numerous studies have, however, attempted to determine the exact relationship between characteristic frequency (CF), F, and distance from the base, x, in the mammalian cochlea. Greenwood published a review of related work in 1990, and determined that a previously derived equation of the form

$$F = A\left[10^{\alpha(35-x)} - k\right], \tag{2.33}$$

with parameters A = 165.4 Hz, α = 0.06 mm^{-1}, and k = 0.88, gave reasonably good results for humans. The frequency-to-place map of the current model was tuned to approximate to this equation, and is shown below in Figure 2.19.

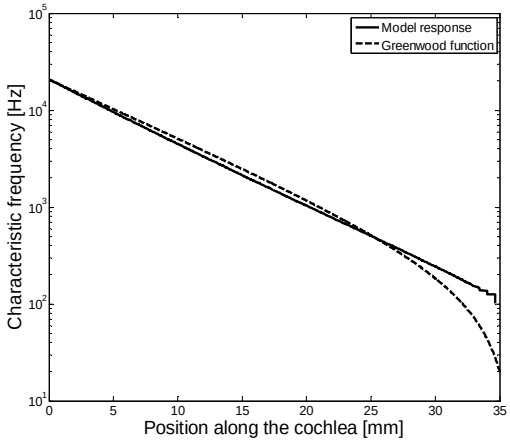


Figure 2.17: Frequency-to-place map of the model and the Greenwood function for a human.

The model does a reasonable job of following the Greenwood curve until $x \approx 29$ mm, where the human tuning curve drops off in frequency more quickly. It was found that this effect could be replicated in the model by adding a larger constant (position independent) amount of damping to $c_1(x)$, but this adversely affected other characteristics of the model and was not applied in the final set of parameters.

Another feature that was deemed important here and has also been measured in a variety of animals is the wavelength of the TW at its peak. By definition,

$$\mathrm{Re}\left[\kappa_{\mathrm{TW}}(x,\omega)\right] = \frac{2\pi}{\lambda_{\mathrm{TW}}(x,\omega)}, \tag{2.34}$$

where λ_{TW} is the wavelength of the TW which varies with both position and frequency. Combining equations (2.17) and (2.34) yields an expression relating the wavelength to cochlear partition impedance:

$$\lambda(x,\omega) = \mathrm{Re}\left[\sqrt{\frac{HZ_{cp}(x,\omega)}{-2j\omega\rho}}\right] 2\pi\,. \tag{2.35}$$

Figure 2.18.a-b shows the variation of the TW wavelength as a function of both position and frequency. As the tonally-generated TW propagates down the cochlea, its wavelength is long in the region where the impedance of the CP is stiffness-dominated. As the local stiffness decreases, so does the TW wavelength. When the wave approaches the resonant position, the wavelength decreases sharply. Although Figure 2.18.a shows that the decrease in λ_{TW} continues apically beyond this point, the magnitude of the TW is strongly

attenuated beyond this position; thus, the TW wavelength is only important from the base up until several mm beyond the characteristic place.

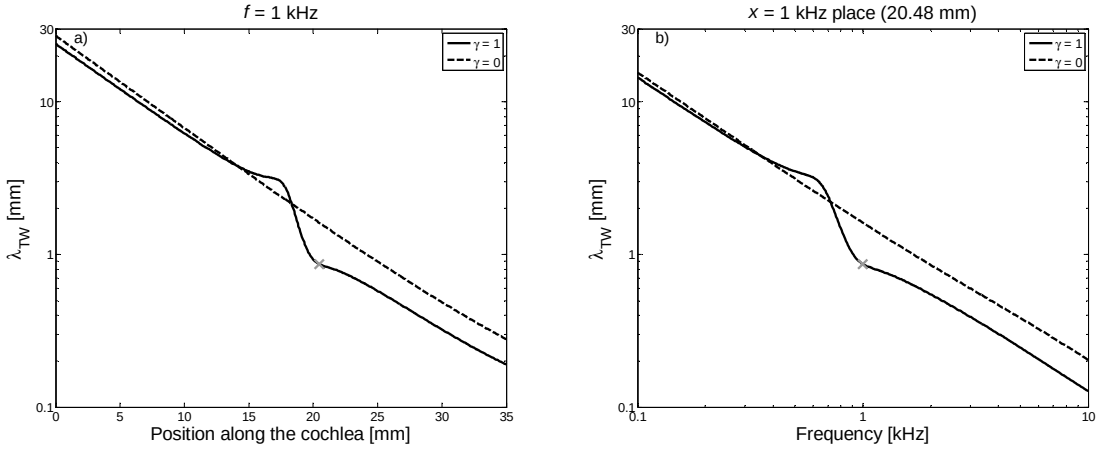


Figure 2.18.a-b: Wavelength of the TW as a function of a) position for a 1 kHz tone, and b) frequency at the 1 kHz characteristic place. A light 'x' denotes the wavelength at the characteristic place (a) and frequency (b) when γ = 1. The active case is shown as a solid line, whereas the passive case is shown as a dashed line.

Similarly, Figure 2.18.b shows that the TW wavelength is long at frequencies lower than the resonant frequency at a given location, as it is basal of those resonant places.

In order to calculate the wavelength of the TW at its peak, λ_{peak}, the CP impedance of equation (2.35) must be evaluated at the characteristic location and frequency:

$$\lambda_{peak}\left(x\right) = \mathrm{Re}\left[\sqrt{\frac{HZ_{cp}\left(x_{cf},\omega_{cf}\right)}{-2j\omega_{cf}\rho}}\right]2\pi, \tag{2.36}$$

where the subscript *cf* denotes characteristic frequency and place. This is plotted in Figure 2.19.

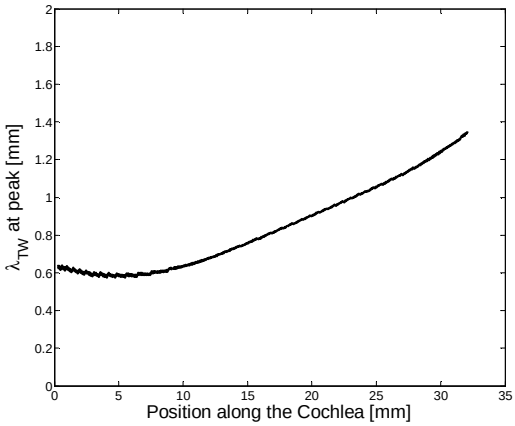


Figure 2.19: Wavelength of the TW at its peak as a function of position along the cochlea.

The data of Figure 2.19 show that the peak wavelength is small at the base of the model and increases roughly linearly toward the apex. This trend is consistent with measurements and extrapolations made in other mammals, though the exact characteristics of the healthy human cochlear TW are as yet unmeasured *in vivo*. However, the values of λ_{peak} in this model are very close to the estimates of Shera and Guinan (2003) based upon their reverse-calculations from OAE data. One issue that Figure 2.19 does raise is that the values of λ_{peak} fall somewhat below 1 mm, the height of the cochlear channel, for much of the model. This presents a problem with the formulation as the longwave assumption is then violated; the implications of this predicament are discussed in Chapter 6.

The imaginary component of the TW wavenumber can provide further information about the system. By analysing the auditory nerve responses of chinchillas and cats, Shera (2007) empirically obtains TW wavenumber characteristics in all regions of physiologically healthy cochleae. The real and imaginary parts of κ_{TW} are referred to as propagation and gain functions in the aforementioned manuscript. These quantities are assigned the Greek letters κ and γ by Shera (2007), but as these are already reserved in this dissertation, they are simply referred to here as the real and imaginary parts of κ_{TW}, $\Re\{\kappa_{\mathrm{TW}}\}$, $\Im\{\kappa_{\mathrm{TW}}\}$. Figure 2.20 illustrates how the propagation and gain functions vary as a function of position at 8 characteristic frequencies in the model, similar to Figure 8 in Shera (2007).

As previously discussed, the real part of the wavenumber is proportional to the inverse of the TW wavelength. Thus, the values of the solid dark lines at the intersections with the dotted vertical lines are equivalent to $2\pi/\lambda_{peak}(x)$. As expected, the wavelength of the TW reaches a minimum at the characteristic place for most frequencies, though some deviation from this trend is observed at lower frequencies which peak in the apex. In addition, the magnitude of the real part of κ_{TW} is somewhat higher in Figure 2.20 than Figure 8 of (Shera, 2007); this implies that λ_{TW} is shorter in humans than in chinchillas. However, when the feedback gain is set to zero, the real part of the TW is much lower in magnitude. This verifies that λ_{TW} varies considerably in the model as a function of γ.

The imaginary part of κ_{TW} is referred to as the 'gain' function in Shera (2007), as it is determines whether the amplitude of the TW is increasing or decreasing. This can be related back to the expression for κ_{TW} in the model. The sign of the real part of the BM

impedance, which is related to the damping, is proportional to the negative product of the real and imaginary components of κ_{TW}, as given in equation (19) of Shera (2007). Thus, because $\Re\left[\kappa_{TW}(\omega)\right] > 0$ for all regions of the cochlear model, the positions where $\Im\left[\kappa_{TW}(\omega)\right] > 0$ correspond to the region of negative damping. This can be seen when comparing Figure 2.20 to the real part of the BM admittance, illustrated in Figure 2.7.b.

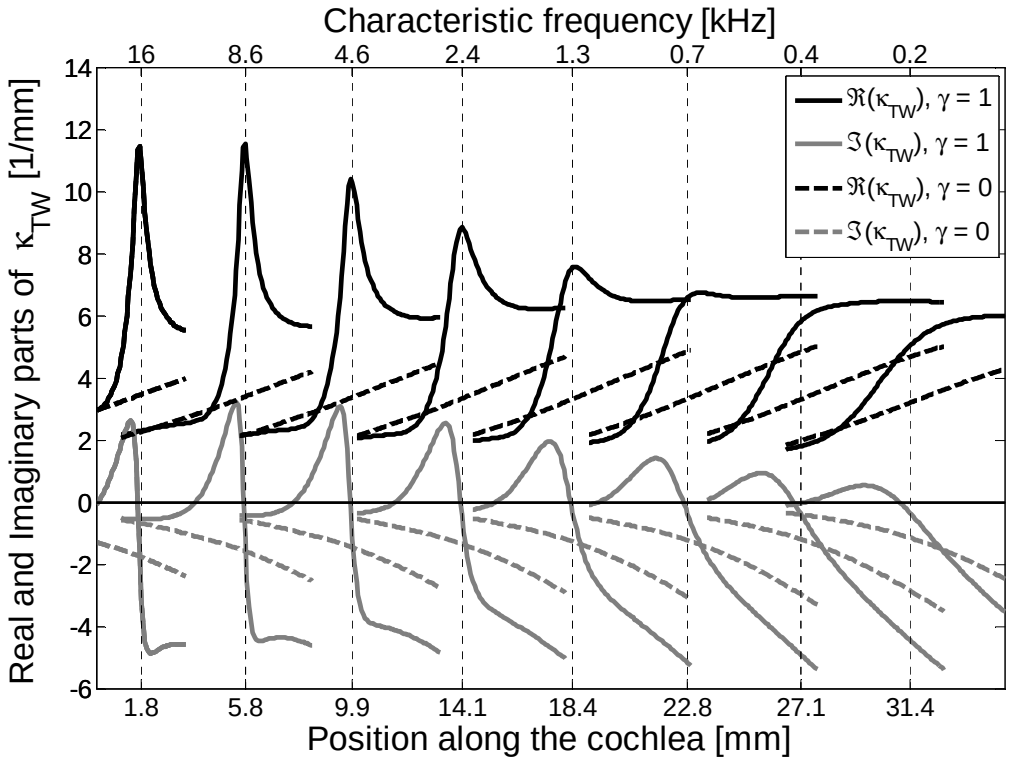


Figure 2.20: Real (dark lines) and imaginary (gray lines) components of the wavenumber for the present model. Solid lines represent the active (γ = 1) results, whereas dashed lines represent the passive (γ = 0) results. Dotted vertical lines mark the characteristic places and frequencies of the stimuli in the coupled cochlea.

Another important quantity that has important consequences for OAE analysis, also related to the wavenumber, is the group delay of the TW, τ_{TW}. The group delay, which is a function of characteristic place and frequency given that the cochlea is a dispersive system, is given by

$$\tau_{TW}\left(x_{cf}, \omega_{cf}\right) = \int_0^{x_{cf}(\omega)} \frac{\partial \,\mathrm{Re}\left(\kappa_{TW}\right)}{\partial \omega} dx, \tag{2.37}$$

(e.g. Moleti *et al*., 2005). Figure 2.21 displays the cochlear group delay as a function of position in the active and passive models. The latency is plot along the horizontal axis in

order to facilitate comparisons between these delays and the propagation of TWs in time domain simulations, as shown in the next chapter.

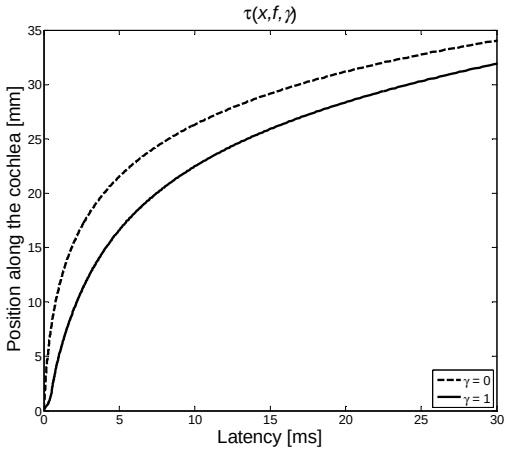


Figure 2.21: Distance to the characteristic place as a function of cochlear delay. The active response is shown as a solid line, whereas the passive response is shown as a dashed line.

The data shows that the group delay is longer when the system is active compared to when it is passive; this is consistent with physiologically measured results (e.g. Recio and Rhode, 2000). Another feature of interest is how many cycles of phase change a tonal TW experiences from the base to its characteristic place.

Figure 2.22 shows the total phase accumulated by the TW between the base and a given characteristic place as a function of frequency; these results are generated by taking the difference between the BM velocity's phase lag at the characteristic place and the phase lead at the base.

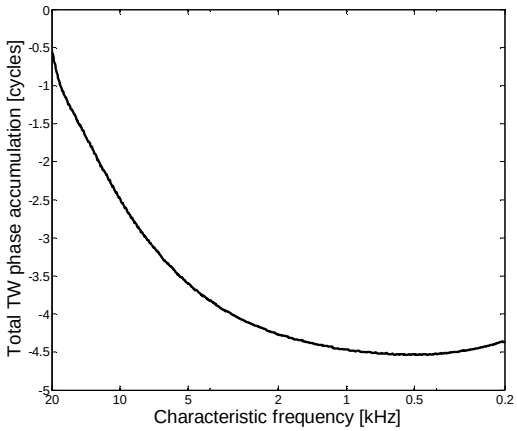


Figure 2.22: Total TW phase accumulation from the base to the characteristic place as a function of characteristic frequency.

Finally, the variation of the model's response is calculated as the number of micromechanical elements in the cochlear model is increased. Although it can be argued that the cochlea is by nature longitudinally discrete, owing to the individual rows of OHCs, the rows are typically ~5 μm apart (Pickles, 2003). This would require dividing the human cochlea up into approximately 7000 segments to create an 'accurate' model by this reasoning. It can be shown, however, that increasing the number of points beyond a certain limit has little practical value. Furthermore, at N = 500, the discretisation size is less than one tenth of a millimetre; this represents a spatial sampling rate that is more than sufficient to meet the Nyquist criterion of this system, given the wavelength of the TW at its peak. Figure 2.23 shows the cochlear response at several frequencies for three different values of N.

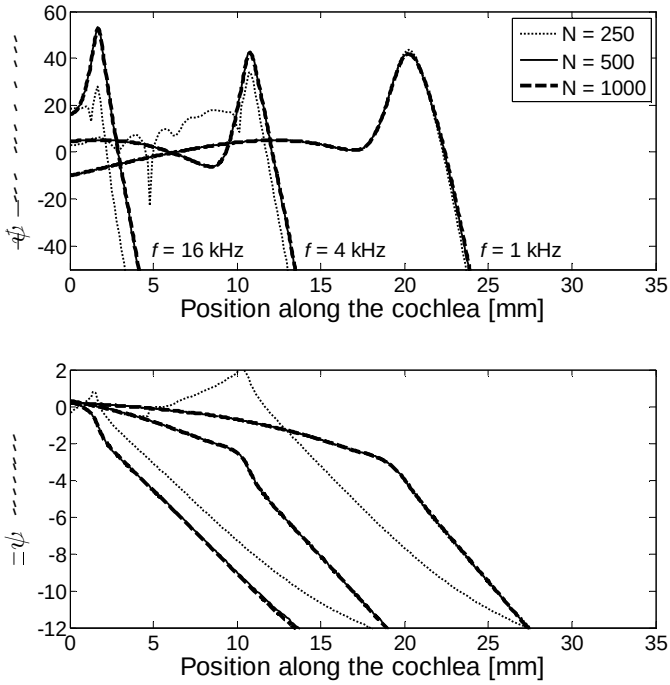


Figure 2.23: Variation of cochlear response given several values of spatial discretisation size.

At N = 250 points, the response exhibits atypical behaviour at higher frequencies. The magnitude and phase are indicative of spurious reflections in the cochlear model that lead to system instability. At N = 500 points, the response is smooth. Doubling N again to 1000 changes the response by less than 1 dB at the peak. However, with N = 250, the

frequency response at 1 kHz looks normal. This potential ambiguity regarding the observed instability of the frequency domain model is one of the reasons why there is a need for a formulation which can unambiguously determine the stability of the fluid-coupled cochlea, as opposed to just that of individual micromechanical elements. This matter is addressed in the following chapter.

Chapter 3

State Space Formulation

Many of the salient features of the biological cochlea, such as its wide dynamic range and the compressive growth of BM motion at moderate stimulus levels, are believed to be partly due to the nonlinearity of the CA (Pickles, 2003). This key feature is omitted in many models because it greatly complicates analysis and nonlinear responses are also often time-consuming to simulate. Toward the beginning of this investigation, a first attempt was made to study the compressive behaviour of the cochlea by implementing Kanis and de Boer's (1993) quasi-linear model.

The quasi-linear approach seeks to approximate nonlinear behaviour in a linear frequency domain model of cochlear mechanics. In a nutshell, this is accomplished through an iterative process that evaluates the linear BM velocity and a compressed OHC pressure in order to generate a quasi-linear CP impedance; this then becomes the basis for the next iteration of linear and compressed results until the responses converge. This procedure has the advantage of computational speed, as all of the calculations are performed in the frequency domain. However, this methodology is still restricted to analysing steady state responses to tonal stimuli, while many interesting features of the cochlea are only transiently expressed. Furthermore, the application of this framework to the Neely and Kim (1986) model generated frequency responses that seemed indicative of reflections and instability. This revealed the need for a rigorous test of a model's stability, as instability in a frequency domain model invalidates its results.

While procedures exist to check the stability of frequency-domain models, these methods often involve calculating the system's frequency response at many different

frequencies to examine its causality. It is difficult to formalise these tests for discretised cochlear models as each individual segment may produce many hundreds of frequency responses. This is of particular concern given that numerical effects, such as truncation in the frequency domain, will also lead to the presence of small non-causal components. The solution adopted in this investigation was to recast the Neely and Kim model in a format known as state space, which is capable of unambiguously determining the stability of the model.

The state space model is introduced in Section 3.1. Its construction is described for both isolated micromechanical elements in Section 3.2, and the fluid-coupled macromechanical system in Section 3.3. The stability of both micro- and macro-mechanical systems is discussed as the feedback gain is increased. The stability is also studied for non-uniform distributions of feedback gain in the coupled model. In Section 3.4, frequency domain simulations generated by the state space model are shown to match the responses of the frequency domain formulation. Finally, both linear and nonlinear time domain simulations of cochlear responses to tones are presented in Section 3.5.

3.1 Formulation overview

The state space approach is inherently set in the time-domain. The dynamics of a system are expressed as a set of coupled first-order differential equations and arranged in vector-matrix form. In the following subsections, the construction of the state space formulation of Neely and Kim's (1986) discretised cochlear model is presented. The standard form of the state equations of a linear time-invariant analog system is given by

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) \end{aligned}, \tag{3.1}$$

where bold letters represent vector-matrices (Phillips and Harbor, 2000). In these equations,

$\mathbf{x}(t)$ is the ($m \times 1$) vector of the states of an mth-order system,

$\mathbf{A}$ is the ($m \times m$) system matrix that contains the mechanics of the model,

$\mathbf{B}$ is the ($m \times r$) input matrix that scales the r input(s) to the system,

$\mathbf{u}(t)$ is the ($r \times 1$) input vector composed of the system input functions,

$\mathbf{y}(t)$ is the ($p \times 1$) output vector composed of the defined p outputs,

C is the ($p \times m$) output matrix that selects the output states of the model, and

D is the ($p \times r$) feed-through matrix that passes the input directly to the output,

where m, r and p are integer values that describe the dimensions of the vector-matrices. Given this format, the stability of the state space system can be determined by calculating the eigenvalues of the system matrix, **A**, as explained in greater detail below.

Subsection 3.2 shows how the cochlear micromechanics can be set in the standard state space form. The stability of the micromechanical system is then investigated in terms of its poles and zeros, as generated by the state space model. Subsection 3.3 then describes how the boundary conditions and fluid-coupling associated with the macromechanical model are set in matrix form and combined with all the micromechanical models, thus yielding the complete state space model of the cochlea.

3.2 Isolated micromechanics

The Neely and Kim (1986) micromechanical model is an active second-order system which has two states associated with each degree of freedom, a single input, and a single output. Thus, for the isolated micromechanical model, $m = 4$, $r = 1$ and $p = 1$. The four state variables are chosen to be the velocity and displacement of the BM and TM:

$$\begin{aligned} \dot{\xi}_p &= \dot{x}_1(t) \\ \xi_p &= x_1(t) \\ \dot{\xi}_t &= \dot{x}_2(t) \\ \xi_t &= x_2(t) \end{aligned} . \tag{3.2}$$

The notation representing BM and TM motion as x_1 and x_2 is adopted here to maintain consistency with standard control theory convention of expressing states as x_i. Note, however, that the overall vector of state variables $\mathbf{x}(t)$ should not be confused with the longitudinal spatial variable x, which is not a function of time. When these four states are vertically concatenated, the state matrix is formed:

$$\mathbf{x}_n = \begin{bmatrix} \dot{x}_1(t) \\ x_1(t) \\ \dot{x}_2(t) \\ x_2(t) \end{bmatrix}_n , \tag{3.3}$$

where the subscript n corresponds to the index number of the element in the cochlea. When the time derivative of equation (3.3) is evaluated, the state matrix becomes

$$\dot{\mathbf{x}}_n = \begin{bmatrix} \ddot{x}_1(t) \\ \dot{x}_1(t) \\ \ddot{x}_2(t) \\ \dot{x}_2(t) \end{bmatrix}_n, \tag{3.4}$$

where $\ddot{x}_1(t)$ and $\ddot{x}_2(t)$ represent the acceleration of the BM and the TM, respectively. Solving the equations of micromechanical motion for the BM and TM acceleration yields

$$\ddot{x}_1 = \frac{1}{m_1}\begin{Bmatrix} p_d(t) + g\gamma\left[c_4\left(\dot{x}_2(t) - \dot{x}_1(t)\right) + k_4\left(x_2(t) - x_1(t)\right)\right] \\ -\dot{x}_1(t)(c_1 + c_3) - x_1(t)(k_1 + k_3) + \dot{x}_2(t)c_3 + x_2(t)k_3 \end{Bmatrix}, \tag{3.5}$$

and

$$\ddot{x}_2 = \frac{1}{m_2}\left\{-\dot{x}_2(t)(c_2 + c_3) - x_2(t)(k_2 + k_3) + \dot{x}_1(t)c_3 + x_1(t)k_3\right\}, \tag{3.6}$$

as derived in Appendix B.4. Recall that the micromechanical parameters vary as a function of position. This notation is suppressed here for convenience. The standard state space form can be rewritten to describe the micromechanics of a given element in the discrete model of the cochlea:

$$\dot{\mathbf{x}}_n(t) = \mathbf{A}_n\mathbf{x}_n(t) + \mathbf{B}_n p_n(t). \tag{3.7}$$

When the components of equations (3.5) and (3.6) are separated into the form of equation (3.7), the micromechanical system matrix, $\mathbf{A}_n$, and input matrix, $\mathbf{B}_n$, are defined as:

$$\mathbf{A}_n = \begin{bmatrix} -\frac{(c_1 + c_3 - g\gamma c_4)}{m_1} & -\frac{(k_1 + k_3 - g\gamma k_4)}{m_1} & \frac{c_3 - \gamma c_4}{m_1} & \frac{k_3 - \gamma k_4}{m_1} \\ 1 & 0 & 0 & 0 \\ \frac{c_3}{m_2} & \frac{k_3}{m_2} & -\frac{(c_2 + c_3)}{m_2} & -\frac{(k_2 + k_3)}{m_2} \\ 0 & 0 & 1 & 0 \end{bmatrix}_n, \tag{3.8}$$

and

$$\mathbf{B}_n = \begin{bmatrix} \frac{1}{m_1} \\ 0 \\ 0 \\ 0 \end{bmatrix}_n. \tag{3.9}$$

The output of the state space model is set as the BM velocity. Rewriting the second line of equation (3.1) yields

$$\dot{\xi}_{p_n}(t) = \dot{x}_{1_n}(t) = \mathbf{C}_n\mathbf{x}_n(t) + \mathbf{D}_n p_d(t), \tag{3.10}$$

where

$$\mathbf{C}_n = \begin{bmatrix} b & 0 & 0 & 0 \end{bmatrix}, \tag{3.11}$$

and

$$\mathbf{D}_n = 0. \tag{3.12}$$

This completes the expression of the micromechanical model in state space form. However, looking forward to the coupled system, it is clear that the middle ear and helicotrema elements must also be expressed in the standard format in order to be compatible with the rest of the model.

The equations of motion for the middle ear and the helicotrema, solved for the acceleration of the element, are given by

$$\ddot{\xi}_m(t) = \frac{1}{m_m}\left[p_1(t) - c_m\dot{\xi}_m(t) - k_m\xi_m(t) \right] \tag{3.13}$$

and

$$\ddot{\xi}_H(t) = \frac{1}{m_H}\left[p_N(t) - c_{st}\dot{\xi}_H(t) \right], \tag{3.14}$$

where $\xi_m(t)$ and $\xi_H(t)$ represent the displacement at the stapes and helicotrema, and $p_1(t)$ and $p_N(t)$ represent the pressure at the stapes and helicotrema. A small amount of damping is added to the helicotrema boundary in order to reduce apical reflections. Following the same procedure as outlined above, the state space matrices for the boundary elements can be expressed as

$$\dot{\mathbf{x}}_1(t) = \mathbf{A}_1\mathbf{x}_1(t) + \mathbf{B}_1 p_1(t), \tag{3.15}$$

and

$$\dot{\mathbf{x}}_N(t) = \mathbf{A}_N\mathbf{x}_N(t) + \mathbf{B}_N p_N(t), \tag{3.16}$$

where

$$\mathbf{x}_1(t) = \begin{bmatrix} \dot{\xi}_m \\ \xi_m \end{bmatrix}, \tag{3.17}$$

$$\mathbf{x}_N(t) = \begin{bmatrix} \dot{\xi}_H \\ \xi_H \end{bmatrix}, \tag{3.18}$$

$$\mathbf{A}_1 = \begin{bmatrix} \dfrac{-c_m}{m_m} & \dfrac{-k_m}{m_m} \\ 1 & 0 \end{bmatrix}, \tag{3.19}$$

$$\mathbf{A}_N = \begin{bmatrix} \dfrac{-c_H}{m_H} & 0 \\ 1 & 0 \end{bmatrix}, \tag{3.20}$$

$$\mathbf{B}_1 = \begin{bmatrix} \frac{1}{m_m} \\ 0 \end{bmatrix}, \tag{3.21}$$

and

$$\mathbf{B}_N = \begin{bmatrix} \frac{1}{m_H} \\ 0 \end{bmatrix}. \tag{3.22}$$

3.2.1 Stability: poles and zeros

It is possible to determine the stability of a model in state space form quickly and unambiguously. A short digression back to the transfer function representation of the micromechanical system will provide a more solid grounding for the discussion of how the state space model accomplishes this. Rearranging and rewriting equation (2.2), the CP admittance, in terms of the Laplace variable, $s = \sigma + j\omega$, gives

$$Y_p(s) = \frac{\dot{\xi}_b(s)}{p_d(s)} = \frac{b\left(Z_2(s) + Z_3(s)\right)}{g\left\{Z_1(s)\left[Z_2(s) + Z_3(s)\right] + Z_2(s)\left[Z_3(s) - \gamma Z_4(s)\right]\right\}}. \tag{3.23}$$

Equation (3.23) is a transfer function of the form

$$H(s) = \frac{N(s)}{D(s)} = K\frac{(s-z_1)(s-z_2)...(s-z_{m-1})(s-z_m)}{(s-p_1)(s-p_2)...(s-p_{m-1})(s-p_m)}, \tag{3.24}$$

where K is the real-valued gain of the system, and the numerator, N(s), and the denominator, D(s), have been factored such that the z_i's are the roots of the equation

$$N(s) = 0, \tag{3.25}$$

and defined to be the system *zeros*, and the p_i's are the roots of the equation

$$D(s) = 0, \tag{3.26}$$

and defined to be the system *poles* (Phillips and Harbor, 2000). The poles and zeros of a system must come in pairs that are either purely real or complex conjugates of each other. When s tends to any of z_i or p_i, the magnitude of the transfer function tends to zero or infinity, respectively:

$$\begin{aligned} \lim_{s \to z_i} H(s) &= 0 \\ \lim_{s \to p_i} H(s) &= \infty \end{aligned}. \tag{3.27}$$

The poles of a system are especially important as they determine the unforced response of a system, $y(t)$; this can be written as a sum of contributions from different modes:

$$y(t) = \sum_{i=1}^{n} C_i e^{p_i t}, \tag{3.28}$$

where the coefficients C_i are dictated by the initial conditions of the system (Levine, 1996).

Defining $\Re\{p_i\} = \sigma_i$ and $\Im\{p_i\} = \omega_i$, it is apparent from equation (3.28) that the envelope of the system response in time will either be stable and decay away if $\sigma_i < 0$, or unstable and grow exponentially if $\sigma_i > 0$. However, poles may also be complex causing the response to oscillate. Thus, poles may contribute to the unforced response of the system in four ways which are illustrated in Figure 3.1.

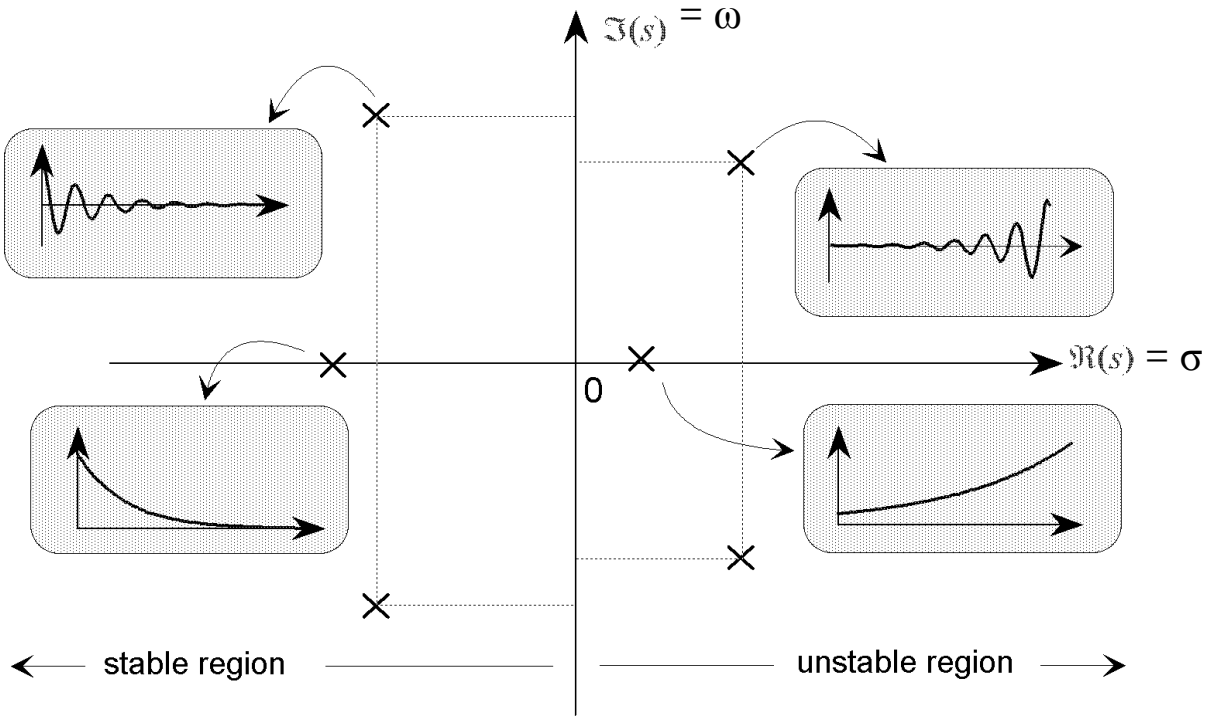


Figure 3.1: Illustration of the effect of poles, marked as 'x's, upon the unforced system response. The real axis is plotted horizontally and the imaginary axis is plotted vertically. Reprinted with permission (Hardt, 2008).

In summary, the positions of the system poles determine whether the system is stable or not.

The poles of the state space model are readily determined by calculating the eigenvalues of the system matrix, **A**. In addition, the zeros can also be determined by solving

$$\det\begin{bmatrix} \mathbf{z}_i\mathbf{I} - \mathbf{A} & -\mathbf{B} \\ -\mathbf{C} & -\mathbf{D} \end{bmatrix} = 0 \tag{3.29}$$

for $\mathbf{z}_i$, (Levine, 1996). Though the system poles dictate the unforced response, the zeros will have a significant impact upon the driven response. The admittance of the Neely and Kim (1986) micromechanical model has 4 poles and 3 zeros. At x = 20.48 mm, there are two complex conjugate poles at $p_{1,2}$ = -856 ± j5790 and two poles on the real axis at p_3 = -7000 and p_4 = -26620, whereas the zeros of this element's admittance are located at $z_{1,2}$ = -800 ± j5000 and z_3 = 0. This is illustrated in Figure 3.2.

The transfer function, $H(s)$ can be solved for any s, but the solution for $s = j\omega$ is of primary interest as this represents the frequency response of the system. The magnitude of the frequency response, as rewritten from equation (3.24), is

$$\left|H\left(j\omega\right)\right| = K\frac{\prod_{i=1}^{m}\left|\left(j\omega - z_i\right)\right|}{\prod_{i=1}^{m}\left|\left(j\omega - p_i\right)\right|}. \tag{3.30}$$

Equation (3.30) is commonly interpreted geometrically. As the driving frequency (ω) is varied, the distances between the s = $j\omega$ point on the imaginary axis and the poles and zeros in the complex s-plane also change. The product of the distances from $j\omega$ to the z_i's divided by the product of the distances from $j\omega$ to the p_i's determines the relative magnitude of the response at different frequencies. An example is given in Figure 3.2 that illustrates the distances from the poles and zeros of Neely and Kim's micromechanical model (at x = 20.48 and γ = 1) to several frequencies on the imaginary axis. Both the passive and active stability and the magnitude of the admittances are plot in Figure 3.3 to motivate this discussion. Note that the position of the zero does not depend on the feedback gain, and almost overlaps with a pole in the passive system.

In Figure 3.2.a, the response is strongly dominated by the zero at the origin. As the frequency is increased, the admittance begins to grow with distance away from the zero at the origin and closer proximity to the poles. At s = j4000, the closest pole or zero is the zero at z = -784 + j4192, which results in a local decrease in admittance. In Figure 3.2.b, the pole at p = -856 + j5790 dominates, and a strong peak is generated. Beyond this frequency, the magnitude begins to decrease. At high frequencies, the fourth pole, located at p = -26,620 + j0 and omitted from panels (a) and (b) for clarity, begins to dominate as the other three relatively closely-spaced poles and zeros cancel each other out. Thus, a frequency of s = j5790*50 only 'sees' this single pole; just as proximity to a pole reinforces the admittance, distance from it diminishes its magnitude.

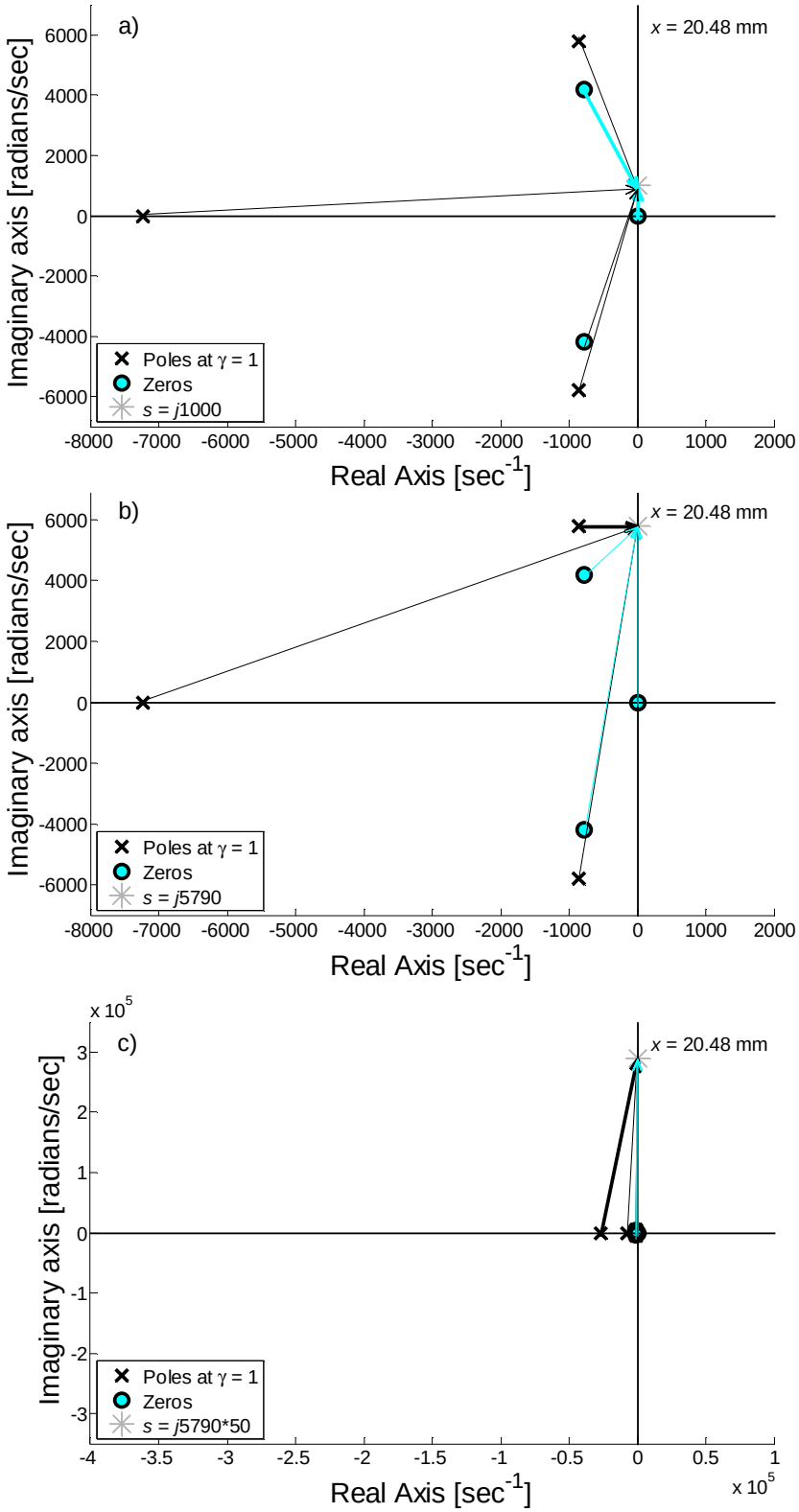


Figure 3.2.a-c: Poles and zeros of Neely and Kim's micromechanical model with $\gamma = 1$. Arrows indicate distances to s, set to 3 frequencies. Light and black lines are drawn from the zeros and poles, respectively. A heavy line denotes the dominant pole or zero in each case. Note that the scale of panel (c) is zoomed out by a factor of 50 relative to (a) and (b).

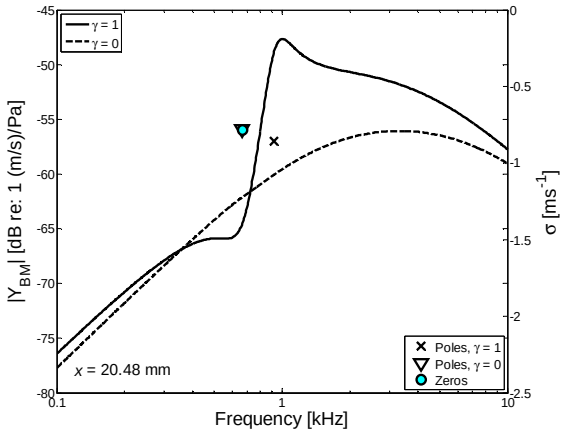


Figure 3.3: Active (—) and passive (--) admittance of the BM at x = 20.48 mm, as in Figure 2.8.a. Also superimposed are the pole and zero in this region. The zero is denoted by a filled circle, whereas poles are plotted as (×) and (▼) given active and passive gains. Note that the axes of the poles and zeros have been rotated, as shown in Figure 3.4.

Figure 3.4 shows how Neely and Kim's (1986) micromechanical pole positions change as the feedback gain, γ, is varied from γ = 0 to γ = 2, at x = 20.48 mm. As predicted by Figure 2.11, this element is unstable at γ > 1.485.

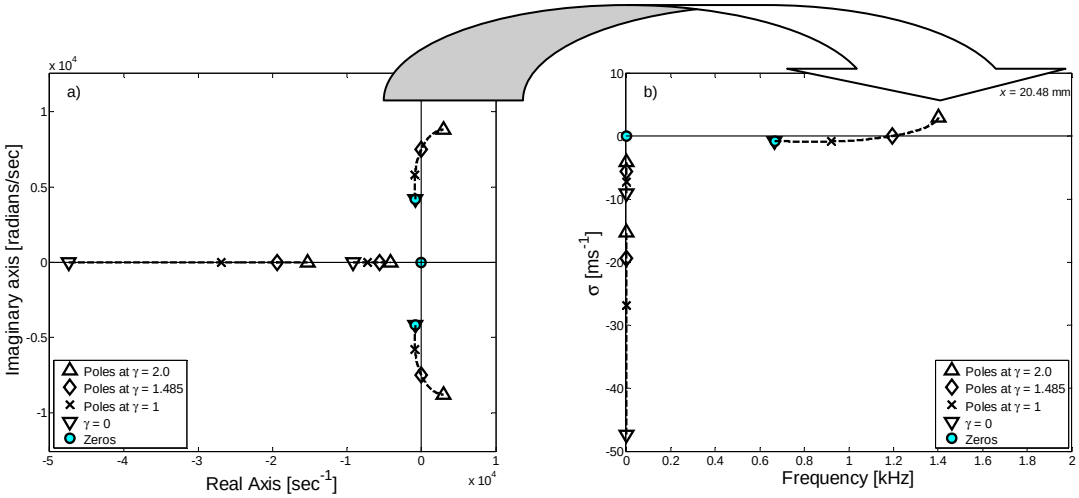


Figure 3.4.a-b: Plot of the zeros and poles of the micromechanical admittance at 20.48 mm along the cochlea given variations in feedback gain, γ. Results are shown in the a) standard form, and b) the adopted 'stability plot' format.

The first thing to notice about Figure 3.4 is that the traditional method of showing poles and zeros is the plot in the (a) panel with, the real axis set horizontally and the imaginary axis set vertically. As the complex poles always appear on the real axis or in conjugate pairs, the negative frequencies can be omitted from the diagram without any loss

of information. Furthermore, the imaginary axis carries units of frequency in radians per second, which does not give much physical insight into the system. For this reason, panel (a) is flipped across the vertical axis, rotated 90 degrees clockwise, and the units along the imaginary axis are converted to kHz. This result is shown in panel (b). This allows direct comparison with frequency responses, as in Figure 3.3. To maintain the consistency of the units between the axes, the real axis is expressed in inverse ms. From this point forward, a plot of poles in the format of Figure 3.4.b is referred to as a 'stability plot,' as poles must fall below the horizontal, $\sigma = 0$, axis for the system to be stable.

The two pairs of poles correspond to the natural frequencies of the passive system, as described in section 2.2. For instance, the poles on the real axis give rise to the heavily-damped higher frequency peak of the BM admittance, whereas the lightly-damped lower frequency peak of the TM admittance gives rise to the conjugate poles of Figure 3.4. At $\gamma = 0$, a complex zero almost exactly overlaps with a complex pole in Figure 3.4.b. Thus, at near-passive gains, the effect of the pole is effectively masked by the zero. This is seen in Figure 3.3, as there is no discernable change in the passive admittance near the frequency of this pole when the system is damped. A small variation is visible in the underamped response of Figure 2.4.a. At $\gamma = 1$ (where the pole is marked by an 'x'), the pole moves steadily away from the zero and increases in frequency. Thus, the magnitude of the response will still be small near the frequency of the zero at non-passive gains, but sharpened near the frequency of the complex pole; this accounts for the dip at 700 Hz and the peak at 1 kHz of the active admittance. Recall that the position of the zero does not depend on the feedback gain, as γ does not appear in the numerator of equation (3.23). As the gain is further increased, the pole moves even closer toward instability and reaches the boundary of stability at $\gamma = 1.485$, which is consistent with Figure 2.11 and Figure 2.12.

As discussed previously, the magnitude of the frequency response will depend on the quotient of the distances between $s = j\omega$ and the zeros and poles. Increasing the gain of the system causes one or more poles to move toward the imaginary axis. When a pole falls on the imaginary axis ($p_i = s = \pm j\omega$), the denominator of the transfer function goes to zero and the response tends to infinity as shown in Figure 2.12. However, when the gain is further increased, the pole begins to move away from the imaginary axis and into the unstable region where $\sigma > 0$; this causes an apparent reduction in the magnitude of the frequency response and a reversal of its sign. This frequency response has, however, now lost

physical validity as the true response in time tends toward infinity. Calculating the pole positions for a linear system is the most reliable method of determining its stability.

The positions of the poles can convey more information than simply if the system is stable or not. For instance, a second-order system can be described by its natural frequency, ω_n, and its damping ratio, ζ; the latter is a dimensionless quantity that describes if and how a system oscillates following an initial perturbation. Geometrically speaking, ζ is the ratio of the negative real component of a pole to its distance from the origin. Thus, the damping ratios of the poles of a system, $p_i = \sigma_i + j\omega_i$, are

$$\zeta = \cos(\alpha) = \frac{-\sigma}{\sqrt{\sigma^2 + \omega^2}}, \tag{3.31}$$

where α is the angle formed between the negative-real half-axis of the s-plane and the pole in question (Phillips and Harbor, 2000). When $\zeta > 1$, the pole pair has two distinct real solutions and the response is referred to as overdamped; this response decays exponentially. When $\zeta = 1$, the pole pair has two identical real solutions and the response is referred to as critically damped. When $0 < \zeta < 1$, the two poles have both real and imaginary parts which are complex conjugates of each other, and the response is referred to as underdamped; this response oscillates and decays away. When $\zeta = 0$, the pole pair has only imaginary components which are complex conjugates of each other and the response is referred to as conditionally stable, as the unforced response rings at the frequency of the poles and neither decays nor grows with time.

In practice, ζ is useful as it allows for the comparison of the response of various poles at different frequencies. Lines that radiate outward from the origin of the stability plot indicate a constant damping ratio. This is exemplified in Figure 3.5 which superimposes the poles from four separate micromechanical models onto the same stability plot. Figure 3.5 indicates that there are a pair of overdamped poles and a pair of underdamped in each micromechanical system. As the underdamped poles have smaller σ_i than the overdamped poles of each location, it is this oscillating and decaying motion that will dominate the unforced response. Panel (b) shows that the damping ratios of the micromechanical system tend to decrease at positions that are closer to the base.

Figure 3.6.a-d shows all the poles of the 500 uncoupled elements on a single stability plot for the passive and the active cases. Although the plot appears to show a continuous line, there are actually many individual poles. These are distinctly visible at the higher

frequencies. The upper 'arm' of underdamped poles has a near-constant damping ratio when the system is passive, as shown in the Figure 3.6.a. When γ is increased to 1, the damping ratios of the underdamped poles begin to vary. ζ is fairly constant at higher frequencies in Figure 3.6.b, but increases as lower frequencies (positions near the apex) are examined. This is due to the relatively stronger contribution of the damping parameters in each element near the apex compared the base. $c_1(x)$, for example, has a constant term that does not vary with position which makes it more dominant near the apex, as the other parameters are still decreasing exponentially. Figure 3.6 includes a single underdamped pole near 1 kHz that corresponds to the middle ear boundary element.

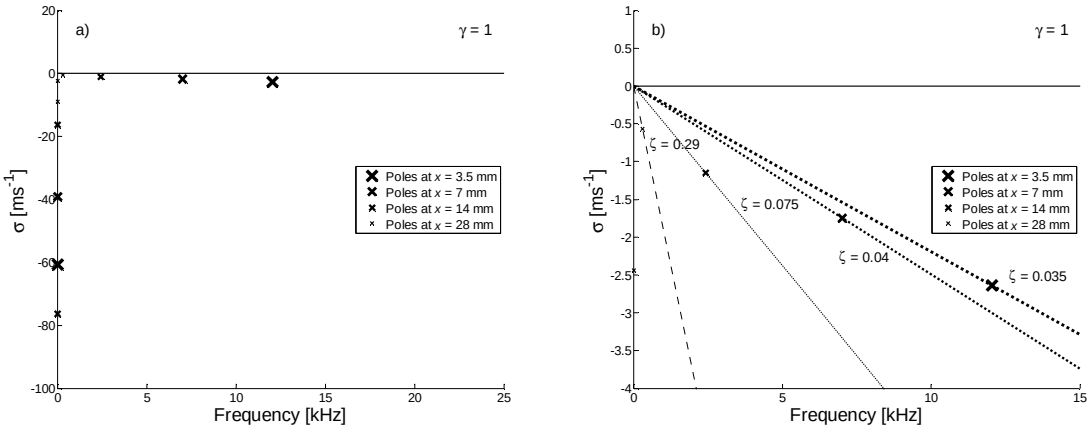


Figure 3.5.a-b: Plot of isolated, active micromechanical poles at four positions along the cochlea. Two more poles at further negative real values beyond the σ-axis limits are omitted for clarity. Differently sized 'x's denote the poles from different positions along the cochlea. b) zoomed-in view of the four underdamped poles near the boundary of stability; dashed lines indicate locations of constant ζ for the underdamped poles.

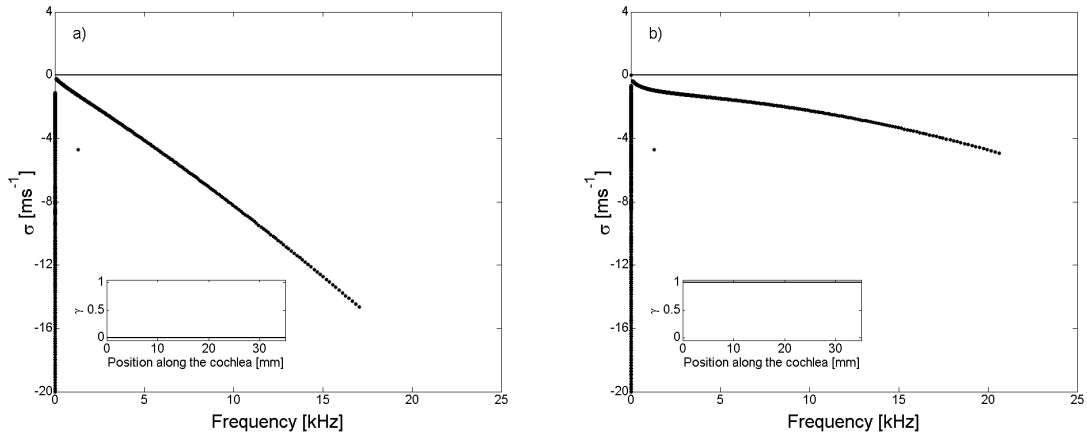


Figure 3.6.a-b: Isolated micromechanical poles plot together at one gain given passive (a) and active (b) models.

3.3 Coupled macromechanics

In order to express Neely and Kim's (1986) coupled macromechanical model in state space form, it is necessary to return to the boundary conditions and the wave equation that underpin this system. The same procedure of discretising the equations presented for the frequency domain model in section 2.4 is carried out here with one notable difference: these equations are now expressed as functions of CP and boundary accelerations in order to comply with the requirements of the state space formulation. Following this, the macromechanics are again set in matrix form. The macromechanical matrices are then combined with the micromechanical matrices set in state space form to generate a coupled state space model of the cochlea.

Rewriting the one-dimensional wave equation of equation (2.16) in terms of the radially averaged acceleration of the cochlear partition yields

$$\frac{\partial^2 p(t)}{\partial x^2} - \frac{2\rho}{H}\ddot{\xi}_p(t) = 0\,, \tag{3.32}$$

where ρ is the density of the cochlear fluids and H is the height of the canal above and below the cochlear partition, which is assumed to be constant. Although both $p(t)$, the pressure difference across the cochlear partition, and $\ddot{\xi}_p(t)$, the radially-averaged acceleration of the cochlear partition, are functions of x, the dependence is suppressed here for notational convenience.

Similarly, the boundary condition for the wave equation at the basal end presented in equation (2.18) can be written in terms of the acceleration at the stapes:

$$\left.\frac{\partial p(t)}{\partial x}\right|_{x=0} = 2\rho\ddot{\xi}_{st}(t)\,, \tag{3.33}$$

where $\ddot{\xi}_{st}(t)$ is the acceleration of the stapes footplate. This has two linearly superposing components: the unloaded part $\ddot{\xi}_{SO}(t)$, which is the acceleration due to an external stimulus, and $\ddot{\xi}_{SR}(t)$, which is the acceleration due to the loading by the internal pressure response in the cochlea at $x = 0$. Thus, equation (3.33) may be written as

$$\left.\frac{\partial p(t)}{\partial x}\right|_{x=0} - 2\rho\ddot{\xi}_{SR}(t) = 2\rho\ddot{\xi}_{SO}(t)\,. \tag{3.34}$$

At the other end of the cochlea, the helicotrema boundary condition can be written as

$$\left.\frac{\partial p(t)}{\partial x}\right|_{x=L} = 2\rho\ddot{\xi}_N' + \frac{1}{H}p_N, \tag{3.35}$$

where $\ddot{\xi}_N'$ is the effective fluid acceleration at the helicotrema. This nonstandard expression allows for a small amount of damping to be added at the apex. The full details and implications of this revision are discussed in Appendix C.2.

As in Chapter 2, finite difference approximations can be applied to discretise the spatial derivatives in equations (3.32), (3.34) and (3.35):

$$\frac{p_{n-1}(t) - 2p_n(t) + p_{n+1}(t)}{\Delta^2} - \frac{2\rho}{H}\ddot{\xi}_n(t) = 0, \tag{3.36}$$

$$\frac{p_2(t) - p_1(t)}{\Delta} - 2\rho\ddot{\xi}_{SR}(t) = 2\rho\ddot{\xi}_{SO}(t), \tag{3.37}$$

and

$$\frac{H}{\Delta^2}\left[\frac{\Delta}{H}p_{N-1}(t) - \left(\frac{\Delta}{H} - \frac{\Delta^2}{H^2}\right)p_N(t)\right] = 2\rho\ddot{\xi}_N', \tag{3.38}$$

respectively. Equations (3.36) through (3.38) can be written in matrix form (Neely, 1981):

$$\mathbf{F}\mathbf{p}(t) - \ddot{\boldsymbol{\xi}}(t) = \mathbf{q}, \tag{3.39}$$

where $\mathbf{p}(t)$ and $\ddot{\boldsymbol{\xi}}(t)$ are the vectors of pressure differences and elemental accelerations, **F** is the finite-difference matrix (also referred to as the 'fluid-coupling' matrix) and $\mathbf{q}$ is the vector of source terms. When the matrices are expanded, equation (3.39) in this case becomes

$$\frac{H}{2\rho\Delta^2}\begin{bmatrix} -\frac{\Delta}{H} & \frac{\Delta}{H} & & & & & & 0 \\ 1 & -2 & 1 & & & & & \\ 0 & 1 & -2 & 1 & & & & \\ & & \ddots & \ddots & \ddots & & & \\ & & & 1 & -2 & 1 & 0 & \\ & & & & 1 & -2 & 1 & \\ 0 & & & & & \frac{\Delta}{H} & -\left(\frac{\Delta}{H}+\frac{\Delta^2}{H^2}\right) & \end{bmatrix}\begin{bmatrix} p_1(t) \\ p_2(t) \\ \\ \vdots \\ \\ p_{N-1}(t) \\ p_N(t) \end{bmatrix} - \begin{bmatrix} \ddot{\xi}_{SR}(t) \\ \ddot{\xi}_2(t) \\ \\ \vdots \\ \\ \ddot{\xi}_{N-1}(t) \\ \ddot{\xi}_N' \end{bmatrix} = \begin{bmatrix} \ddot{\xi}_{SO}(t) \\ 0 \\ \\ \vdots \\ \\ 0 \\ 0 \end{bmatrix}. \tag{3.40}$$

In order to combine the macromechanics above with the state space representation of the micromechanics presented in the previous subsection, the isolated elements must be collected in the combined matrix equations:

$$\dot{\mathbf{x}}(t) = \mathbf{A}_E\mathbf{x}(t) + \mathbf{B}_E\mathbf{p}(t), \tag{3.41}$$

and

$$\dot{\boldsymbol{\xi}}_p(t) = \mathbf{C}_E \mathbf{x}(t), \tag{3.42}$$

where the vectors are defined as

$$\mathbf{x}^T(t) = \begin{bmatrix} \mathbf{x}_1^T(t) & \mathbf{x}_2^T(t) & \cdots & \mathbf{x}_{N-1}^T(t) & \mathbf{x}_N^T(t) \end{bmatrix}, \tag{3.43}$$

$$\dot{\boldsymbol{\xi}}_p^T(t) = \begin{bmatrix} \dot{\xi}_1(t) & \dot{\xi}_2(t) & \cdots & \dot{\xi}_{N-1}(t) & \dot{\xi}_N(t) \end{bmatrix}, \tag{3.44}$$

and

$$\mathbf{p}^T(t) = \begin{bmatrix} p_1(t) & p_2(t) & \cdots & p_{N-1}(t) & p_N(t) \end{bmatrix}, \tag{3.45}$$

where the superscript T denotes the transpose of the matrix. The elemental matrices, which are block diagonal, are defined as

$$\mathbf{A}_E = \begin{bmatrix} \mathbf{A}_1 & 0 & \cdots & & \\ 0 & \mathbf{A}_2 & & & \\ \vdots & & \ddots & & \vdots \\ & & & \mathbf{A}_{N-1} & 0 \\ & & \cdots & 0 & \mathbf{A}_N \end{bmatrix}, \tag{3.46}$$

$$\mathbf{B}_E = \begin{bmatrix} \mathbf{B}_1 & 0 & \cdots & & \\ 0 & \mathbf{B}_2 & & & \\ \vdots & & \ddots & & \vdots \\ & & & \mathbf{B}_{N-1} & 0 \\ & & \cdots & 0 & \mathbf{B}_N \end{bmatrix}, \tag{3.47}$$

and

$$\mathbf{C}_E = \begin{bmatrix} \mathbf{C}_1 & 0 & \cdots & & \\ 0 & \mathbf{C}_2 & & & \\ \vdots & & \ddots & & \vdots \\ & & & \mathbf{C}_{N-1} & 0 \\ & & \cdots & 0 & \mathbf{C}_N \end{bmatrix}. \tag{3.48}$$

Solving the macromechanical equation (3.39) for the pressure matrix, **p**, gives

$$\mathbf{p}(t) = \mathbf{F}^{-1}\ddot{\boldsymbol{\xi}}(t) + \mathbf{F}^{-1}\mathbf{q}(t), \tag{3.49}$$

where $\mathbf{F}^{-1}$ is the inverse of the fluid-coupling matrix, **F**. Note from equation (3.42) that $\ddot{\boldsymbol{\xi}}_p(t)$ is equal to $\mathbf{C}_E\dot{\mathbf{x}}(t)$, so that

$$\mathbf{p}(t) = \mathbf{F}^{-1}\mathbf{C}_E\dot{\mathbf{x}}(t) + \mathbf{F}^{-1}\mathbf{q}(t). \tag{3.50}$$

Substituting this result into equation (3.41) allows the fluid-coupled cochlea with distributed micromechanics and dynamic boundary conditions to be written in the general state space form,

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t), \tag{3.51}$$

where

$$\mathbf{A} = \left[\mathbf{I} - \mathbf{B}_{\mathrm{E}}\mathbf{F}^{-1}\mathbf{C}_{\mathrm{E}}\right]^{-1}\mathbf{A}_{\mathrm{E}}, \tag{3.52}$$

$$\mathbf{B} = \left[\mathbf{I} - \mathbf{B}_{\mathrm{E}}\mathbf{F}^{-1}\mathbf{C}_{\mathrm{E}}\right]^{-1}\mathbf{B}_{\mathrm{E}}, \tag{3.53}$$

and

$$\mathbf{u}(t) = \mathbf{F}^{-1}\mathbf{q}(t). \tag{3.54}$$

Off-diagonal terms are generated in the system matrix, **A**, when the uncoupled $\mathbf{A_E}$ matrix is multiplied by the quantity $\left[\mathbf{I} - \mathbf{B}_{\mathrm{E}}\mathbf{F}^{-1}\mathbf{C}_{\mathrm{E}}\right]^{-1}$ in equation (3.52); this effectively allows the previously isolated micromechanical models to affect the dynamics of nearby elements.

3.3.1 Stability: uniform variation of γ(*x*)

The stability of the coupled cochlea is only briefly examined in the following subsections. Much greater attention is given to the underlying causes and characteristics of instability in the next chapter, which is concerned with spontaneous emissions.

The poles of the coupled model are determined in the exact same manner as those of the isolated micromechanical model in state space: by calculating the eigenvalues of the system matrix, **A**. While simple to compute in state space, the poles are much less straightforward to calculate from the coupled frequency domain model, hence the need for this formulation. Conversely, the zeros of the state space model, as given by (3.29), are extremely time-consuming to determine when the state matrices are large and are omitted from the discussion here. Figure 3.7 shows the change in pole positions as the feedback gain is uniformly varied as a function of position.

Going from the passive model, shown panel (a), to the active model, shown in panel (b), of Figure 3.7, shows an upward stretch in frequencies of the upper 'arm' of poles, from a maximum of approximately 16 kHz to 20 kHz. As with the isolated system, the upper arm of poles moves upward toward the boundary of stability as the gain is increased. However, the system now becomes unstable at $\gamma = 1.06$, which is less than the lowest value of minimum unstable gain for the isolated micromechanical system ($\gamma = 1.18$), as shown in Figure 2.11. As expected, the system becomes more unstable and further un-damped across a wider range of frequencies as the gain is further increased (Figure 3.7.d); the coupled cochlear model is clearly unstable at $\gamma = 1.20$.

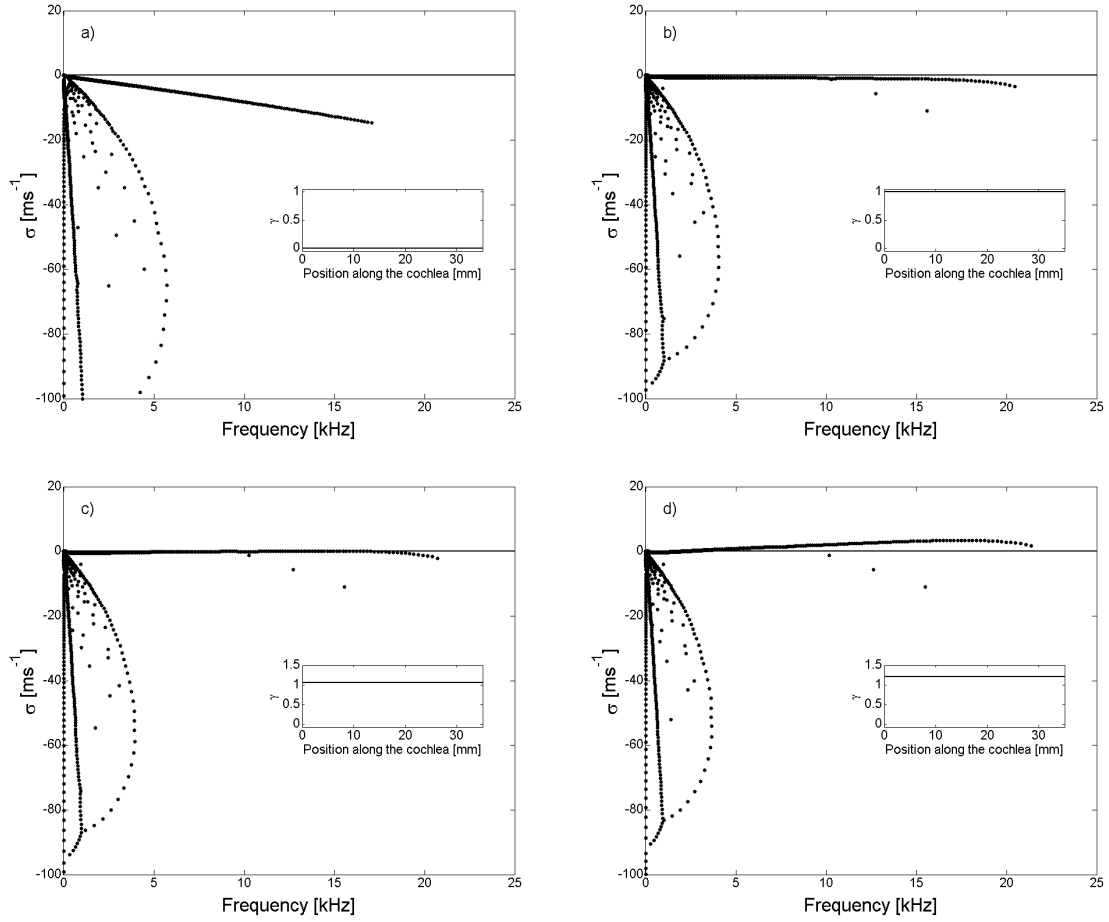


Figure 3.7.a-d: Stability plots of the coupled state space model: a) $\gamma = 0$, system is stable; b) $\gamma = 1$, system is stable; c) $\gamma = 1.06$, system is unstable; d) $\gamma = 1.20$, system is unstable.

3.3.2 Stability: step change in γ(*x*)

It has been previously reported that non-smooth variations in CP impedance as a function of longitudinal position can result in instability. Although such inhomogeneities could result from variations in any of the physical parameters in the micromechanical model, the feedback gain is perturbed here to illustrate the effects on stability. A step change in the feedback gain as a function of position represents a simple spatial discontinuity; the stability of the coupled model given a step change from $\gamma = 1$ to $\gamma = 0.9$ and vice versa at approximately 11 mm is presented in Figure 3.8.a-d.

Figure 3.8 shows that the system is unstable for a 10% step decrease in gain at 11 mm, but remains stable for a step increase in gain, from $\gamma = 0.9$ to $\gamma = 1$, at the same location. The lower panels, Figure 3.8.c-d, show enlarged versions of the upper graphs at

approximately 4 kHz. Examining the upper arms of poles in Figure 3.8.b, it is clear that the poles at frequencies above 4 kHz (the CF at the location of the discontinuity in gain) are more damped than those of the same region in Figure 3.8.a; this is consistent with the distributions of gain in the two models. It is also notable that there is a seemingly-regular spacing between the poles that move toward instability, as shown most clearly in panels (c) and (d) of Figure 3.8.

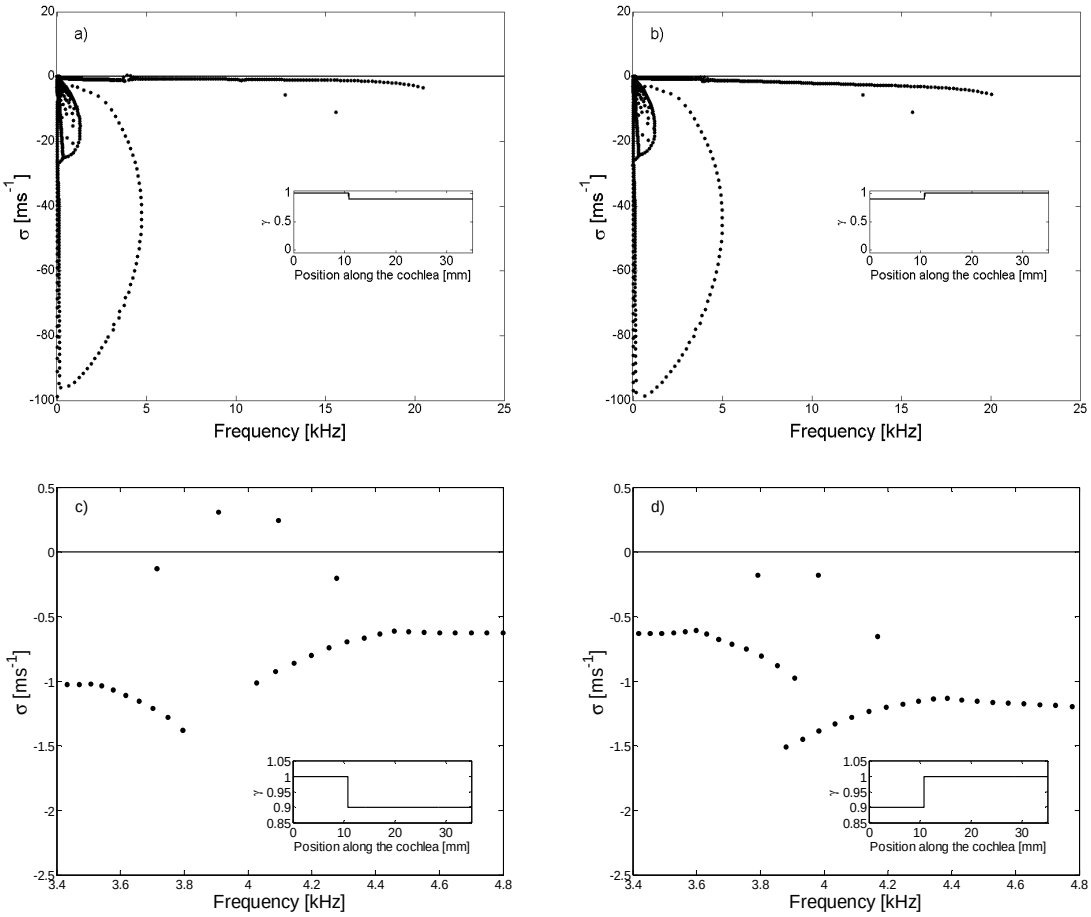


Figure 3.8.a-d: Stability of two coupled cochlear models given step changes in γ. a) γ($x \leq$ 11 mm) = 1, γ(x > 11 mm) = 0.9, system is unstable; and b) γ($x \leq$ 11 mm) = 0.9, γ(x > 11 mm) = 1, system is stable. Panels c) and d) show zoomed-in plots of a) and b), respectively.

3.3.3 Stability: random variation of γ(x)

In most biological systems, there will be a degree of developmental randomness in the structures (e.g. Lonsbury-Martin *et al.*, 1987). In the case of the cochlea, the regularity in the geometric patterns of OHC stereocilia may be imperfect, as shown by Lonsbury-Martin *et al*. (1987) who studied cochleae of rhesus-monkeys. The quantitative variability that

exists in a human cochlea is unknown, so a range of perturbations are applied. Inhomogeneities are applied to the smooth spatial variation of feedback gain with band-passed random waveforms exhibiting a Gaussian distribution. A 5^{th} order Butterworth filter was chosen for its characteristically flat passband (Lineton, 2001).

Figure 3.9.a-d shows the stability of two models, each with a different distribution of feedback gains; panels c-d show the upper arm of poles in detail. The low corner wavenumber of the bandpass filter used to determine the random spatial variations was fixed at 35 mm, the length of the cochlea, in order to avoid introducing large DC shifts in the gain. The high corner wavenumber was set to a different value for each model, one close to the spatial discretisation limit thus generating a 'rough' distribution, and one closer to the low corner wavenumber thus generating a 'smooth' distribution. A more rigorous definition of 'rough' and 'smooth' distributions is introduced in the next chapter.

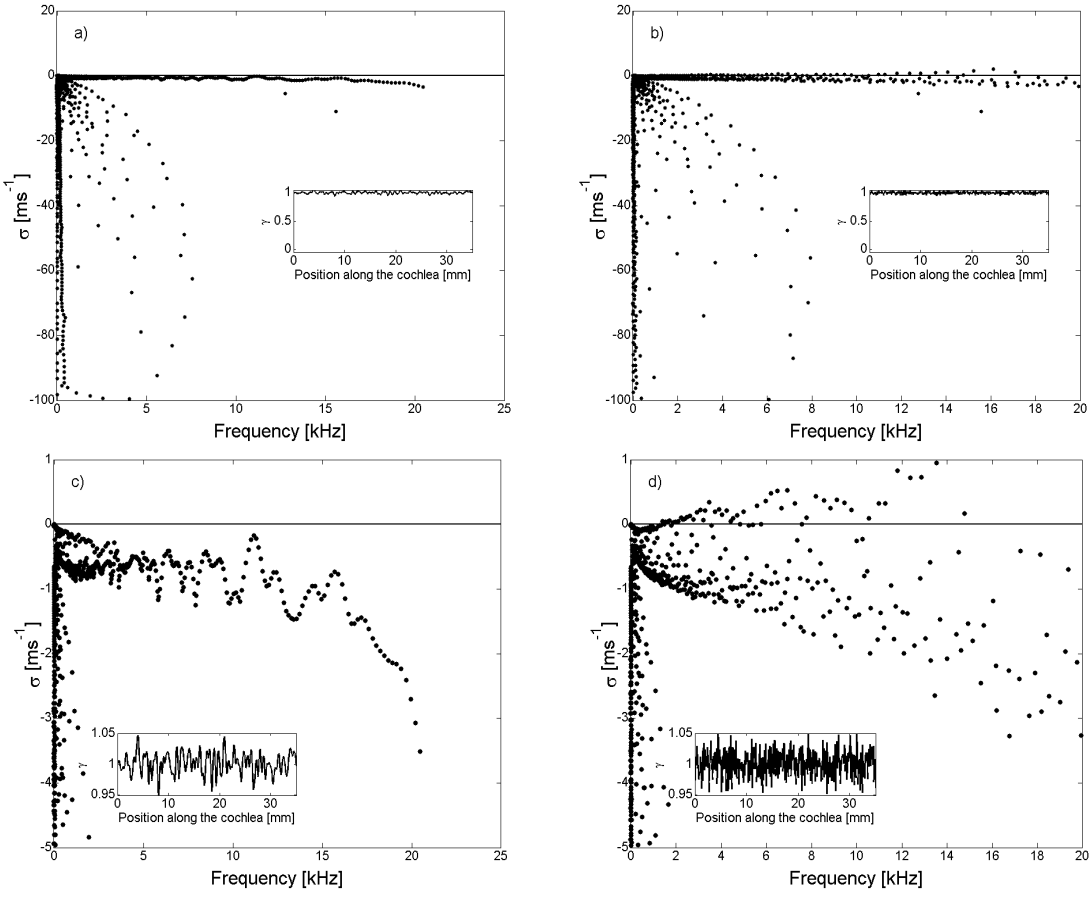


Figure 3.9.a-d: Stability of two models given smooth (a, c) and rough (b, d) distributions of feedback gain. Panels (c) and (d) show a zoomed-in view of panels (a) and (b), respectively.

While both rough and smooth perturbations of gain have a ±5% peak-to-peak deviation from $\gamma = 1$, only the rough distribution generates instability in the model. Panel (d) appears to show two distinct, linear trends of poles away from the origin—one with a stable (positive) damping ratio and one with an unstable (negative) damping ratio. These characteristics are more easily interpreted given a framework of the underlying generation mechanisms of instabilities in the cochlea, as presented in the following chapter.

3.4 Frequency domain responses

It is possible to generate frequency domain responses from the state space model, as outlined below. This is not generally how frequency responses are computed as it is more computationally intensive than simply using the frequency domain model introduced in Section 2.4. Nevertheless, it is a good check to compare results between the two formulations.

In order to find a frequency domain expression for the states of the model, it is necessary to take the Laplace transform of the general state space equations given in (3.1):

$$\begin{aligned} s\mathbf{X}(s) &= \mathbf{A}\mathbf{X}(s) + \mathbf{B}\mathbf{U}(s) \\ \mathbf{Y}(s) &= \mathbf{C}\mathbf{X}(s) + \mathbf{D}\mathbf{U}(s) \end{aligned}. \tag{3.55}$$

The first line is then solved for X(s):

$$\mathbf{X}(s) = (s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}\mathbf{U}(s), \tag{3.56}$$

which can be substituted into the second line of (3.55) yielding

$$\mathbf{Y}(s) = \left[\mathbf{D} + \mathbf{C}(s\mathbf{I} - \mathbf{A})\mathbf{B}\right]\mathbf{U}(s), \tag{3.57}$$

where in this case the output is the BM velocity and $\mathbf{D} = 0$. Taking the Laplace transform of equation (3.54) gives

$$\mathbf{U}(s) = \mathbf{F}^{-1}\mathbf{Q}(s). \tag{3.58}$$

Similarly, to determine the pressure response in the frequency domain, the Laplace transform is taken of equation (3.50) to give

$$\mathbf{P}(s) = \mathbf{F}^{-1}\mathbf{C}s\mathbf{X}(s) + \mathbf{F}^{-1}\mathbf{Q}(s), \tag{3.59}$$

where **X**(s) can be calculated as a function of **Q**(s) with equations (3.56) and (3.58). The frequency response can then be obtained by setting $s = j\omega$, provided the system is stable. Figure 3.10 compares the BM velocity and pressure difference frequency response

calculated within this framework against that of the frequency domain formulation at a stimulus frequency of 4 kHz, given a constant stapes velocity of 1 m/s. The response of the state space model matches that of the frequency domain model as a function of position up to the limit of the state space formulation's dynamic range; this limit is thought to be due to the numerical issues associated with the inversion of the **F** matrix. In the case of the BM velocity, shown in the (b) panel, the magnitude of the response breaks down approximately 260 dB below the peak of the response.

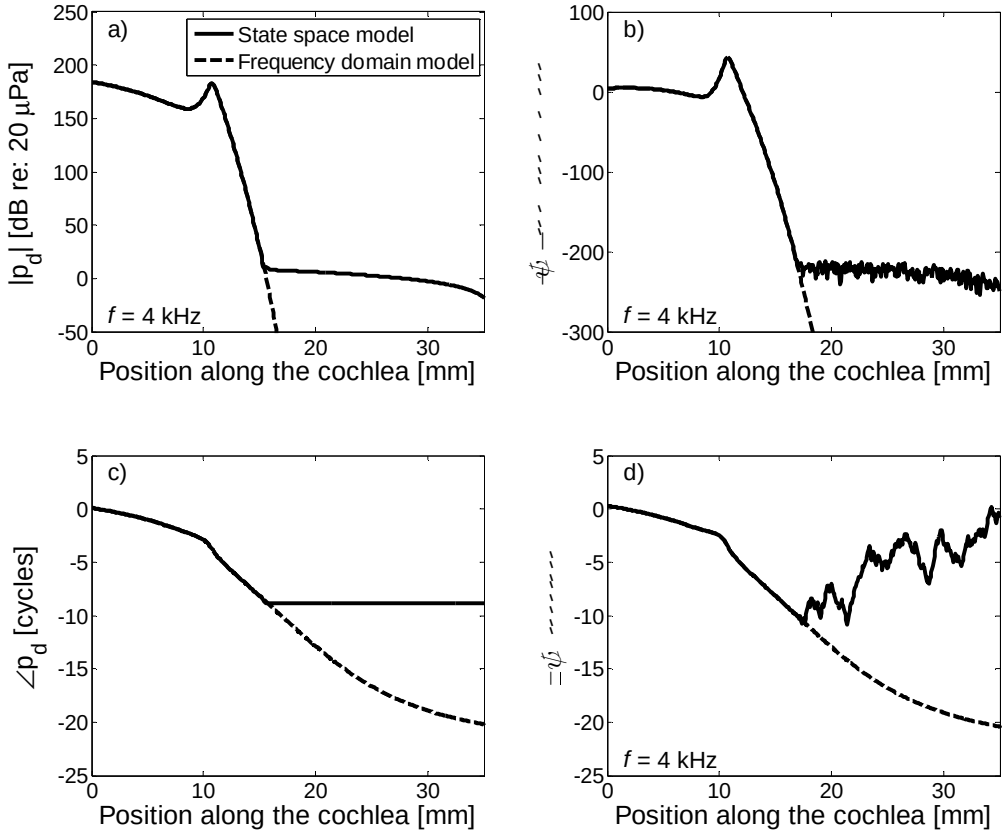


Figure 3.10: Response of the model calculated using the state space formulation at 4 kHz (solid) compared with the frequency domain formulation (dashed). The (a) and (c) panels show magnitude and phase of the pressure difference along the BM, while the (b) and (d) panels show the magnitude and phase of the BM velocity.

3.4.1 Input and output impedances

It is also useful to calculate the input and output impedances of the cochlea for comparison against experimental measurements. Z_{in} is defined as the impedance looking in toward the

cochlea from the middle ear, whereas Z_{out} is defined as the impedance looking out toward the middle ear from the cochlea.

In order to calculate Z_{in} numerically, the middle ear is removed from the model and the output pressure at the first element of the cochlea (P_{st} at $x = 0$ in the coupled cochlea) is divided by the input volume velocity, u_{stapes}, times the area of the stapes footplate, A_s:

$$Z_{\text{in}} = \frac{P_{st}}{u_{stapes} * A_s}. \tag{3.60}$$

However, the Neely and Kim model does not account for the mechanical stiffness of the round window; this is believed to be responsible for the unexpected phase lead of Z_{in} otherwise observed at low frequencies, as shown in Figure 3.11.b. In theory, inclusion of a round window stiffness term should result in an impedance that has the form:

$$Z'_{in} = Z_{\text{in}} + \frac{k_{rw}}{j\omega}. \tag{3.61}$$

Z_{out} is simply the impedance of the middle ear,

$$Z_{out}(j\omega) = j\omega M_m + C_m + \frac{K_m}{j\omega}. \tag{3.62}$$

An empirically derived stiffness of $K_{rw} = 4*10^9$ [N*m^{-3}] gives the correct phase shift at low frequencies without significantly altering the magnitude of Z_{in}.

The magnitude and phase of Z_{in} and Z_{out} are plot side-by-side in Figure 3.11 with physiologically measured data supplied by Professor Sunil Puria, as published in Puria (2003). The magnitude and phase of Z_{out} fall within 1 standard deviation of measured data for most of the frequency range of interest. The magnitude of Z_{in} is on the same order as that of measured data, and its phase response approaches a lag at lower driving frequencies with the inclusion of k_{rw}.

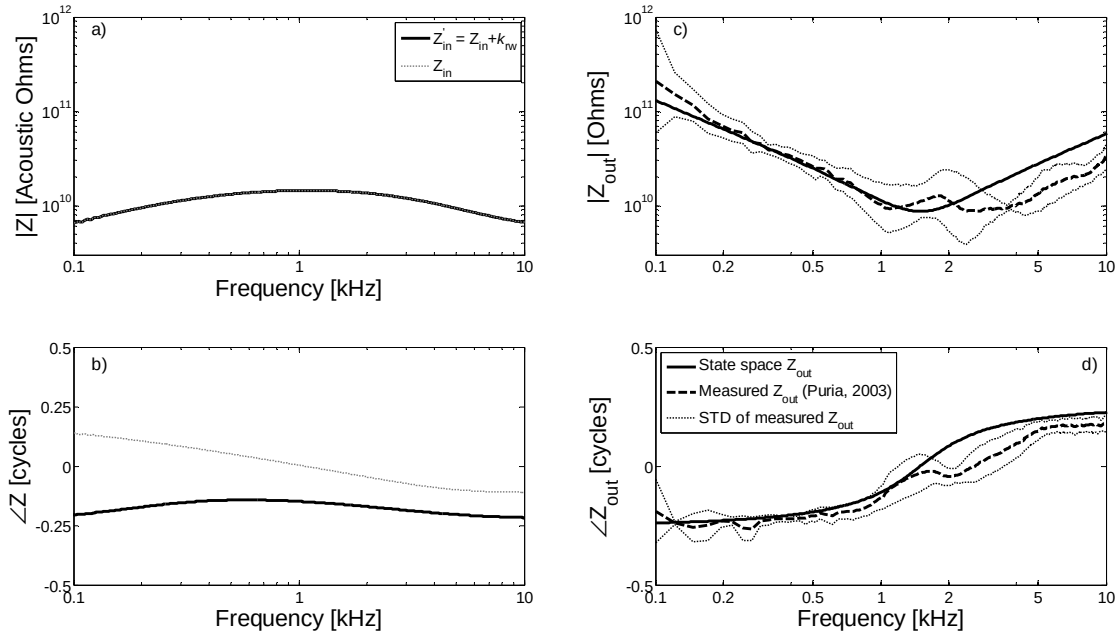


Figure 3.11.a-d: Comparison of Z'_{in} (a, c) and Z_{out} (c, d) magnitude (a, c) and phase (b, d). The light, dashed line in the (a, b) panels shows Z_{in} without the added round window stiffness ($k_{rw} = 4*10^8$ [N*m^{-3}]). Experimentally measured values are plot along with model data in all panels (Puria, 2003).

One other characteristic of cochlear models that has been reported to have important consequences on the stability of the system is the magnitude of the reflection coefficient at the basal boundary (Shera and Zweig, 1991a). The reflection coefficient at the middle ear boundary, as driven from within the cochlea, is defined as the quotient of the forward- and backward- travelling pressure waves, or equivalently, forward- and backward- wave impedances. However, the definition of the forward and backward components, or basis waves, is critical to this analysis. Preliminary work in this area has utilised the WKB method to calculate an approximation of the local characteristic impedance of the cochlea at the base (de Boer and Viergever, 1983; Viergever and de Boer, 1986), though other work has suggested that applying Hankel functions may be more appropriate (Shera and Zweig, 1991a-b). In addition, it should be possible to calculate the reflection coefficient numerically as well. This is an important area of work that deserves further attention and is discussed in Chapter 6.

3.5 Time domain simulations

Time domain simulations can give greater insight into the mechanics of the cochlea as a wide variety of stimuli and analytical techniques can be applied. The transient activity of the model is obscured when only steady state responses are generated in the frequency domain. Crucially, time domain simulations also allow for nonlinearities to be incorporated into the model. The cost of these advantages is the computationally intensive nature of these simulations, relative to frequency domain simulations.

In this subsection, the time-domain response of the state space model is studied. This is not an exhaustive study of all possible simulations that can be performed; rather, it is an introduction to the time domain. Only tonal responses generated in baseline, stable cochleae are considered here. However, both linear and nonlinear responses are calculated; the latter shows harmonic distortion at moderate driving levels. The simulation of transient stimuli and unstable or otherwise perturbed models is left for later chapters.

3.5.1 Linear responses

All time domain simulations are accomplished using MATLAB's ordinary differential equation solver, *ode45*. Internal to this function, the time step is adjusted both forward and backward until the solution settles within supplied error tolerance limits. Absolute and relative tolerances are typically set to 10^{-12} and 10^{-10}, respectively, though these may be adjusted upward and downward depending on the amplitude of the stimuli. For instance, a 0 dB signal would require an absolute error tolerance 100 times smaller than a 40 dB signal. The data is output at a given sampling rate specified by the input time vector; this is set to 50 kHz to avoid any potential aliasing at the highest characteristic frequencies.

In order to reduce the amount of time required for tonal stimuli to reach steady state, a half-Hanning window is applied to the first 10 ms of the stapes acceleration that is the input to the simulation. This reduces the amplitude of the initial impulse that propagates through the cochlea due to the sudden onset of the tone. Figure 3.12 shows the abruptly applied stimulus (stapes acceleration, as generated by a 0 dB SPL 3 kHz tone suddenly input to the ear canal, as detailed in Appendix A), the half-Hanning window function and the windowed stimulus.

A time domain simulation of the 3 kHz stimulus tone shown in Figure 3.12 was set to run for 100 ms. In order to extract the steady-state magnitude and phase information,

only the last 30 ms (90 cycles) of data are analysed. The envelope of the response was extracted by evaluating the maximum values of BM velocity as a function of position across time. Qualitatively similar results were calculated when the RMS amplitude as a function of position was calculated; however, the fine structure of the response is better resolved by taking the maximum value. The phase of the response was calculated by evaluating the discrete Fourier transform (DFT) at 3 kHz at each position along the cochlea. This data is plot along with the frequency domain result generated by the state space model in Figure 3.13.

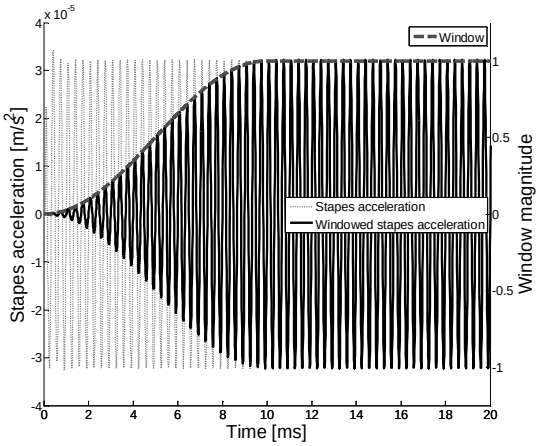


Figure 3.12: The first 20 ms of stapes acceleration that serves as an input to the time domain simulation. The stapes acceleration (dotted, light) is scaled by a half-Hanning window of duration 10 ms (dashed, thick). The resultant windowed stapes acceleration is presented as a solid black line.

The time domain simulation required approximately 8 hours to complete on a desktop computer with a 3.4 GHz Pentium 4 processor and 2 gigabytes of RAM when tight absolute and relative error tolerances of 10^{-15} and 10^{-13} were set, respectively. In contrast, the state space frequency domain simulation required approximately 18 seconds to complete. For comparison, the frequency domain model described in Chapter 2 can compute the response at a single frequency in approximately 43 ms. The noise floor of the frequency domain response is approximately 150 dB below that of the time domain simulation in Figure 3.13.

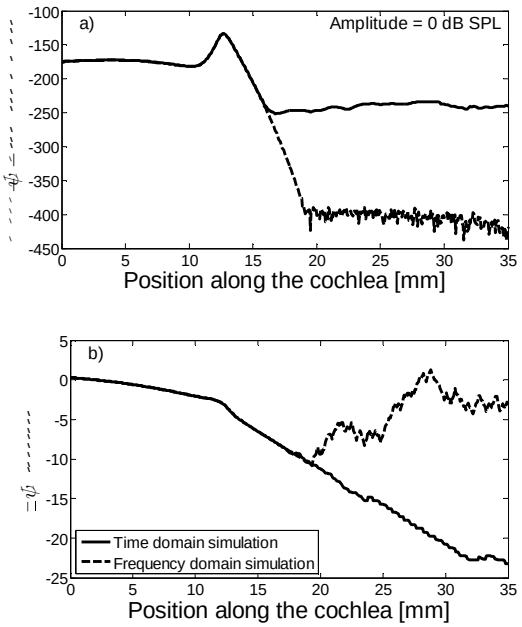


Figure 3.13: Magnitude and phase of a 3 kHz stimulus tone plotted as a function of position along the cochlea. Results are obtained from a linear time domain simulation (solid line) and a state space frequency domain simulation (dashed).

However, frequency domain results cannot show the evolution of the response in time. The first 30 ms of the 3 kHz BM velocity response along the cochlea are shown in Figure 3.14 as a mesh plot.

There appears to be no activity in the first 4 ms or so of Figure 3.14; this is simply due to the fact that the amplitude of the response is much smaller at this time frame when compared to later time frames where the stimulus is no longer windowed. The positive 'slope' of the undulations on the mesh indicates that the TW is propagating forward with increasing time. Note, however, how the slope becomes much less steep near the peak of the response; this indicates that the TW is much slower in the region of the peak, as expected. Beyond this point spatially, the TW is quickly extinguished.

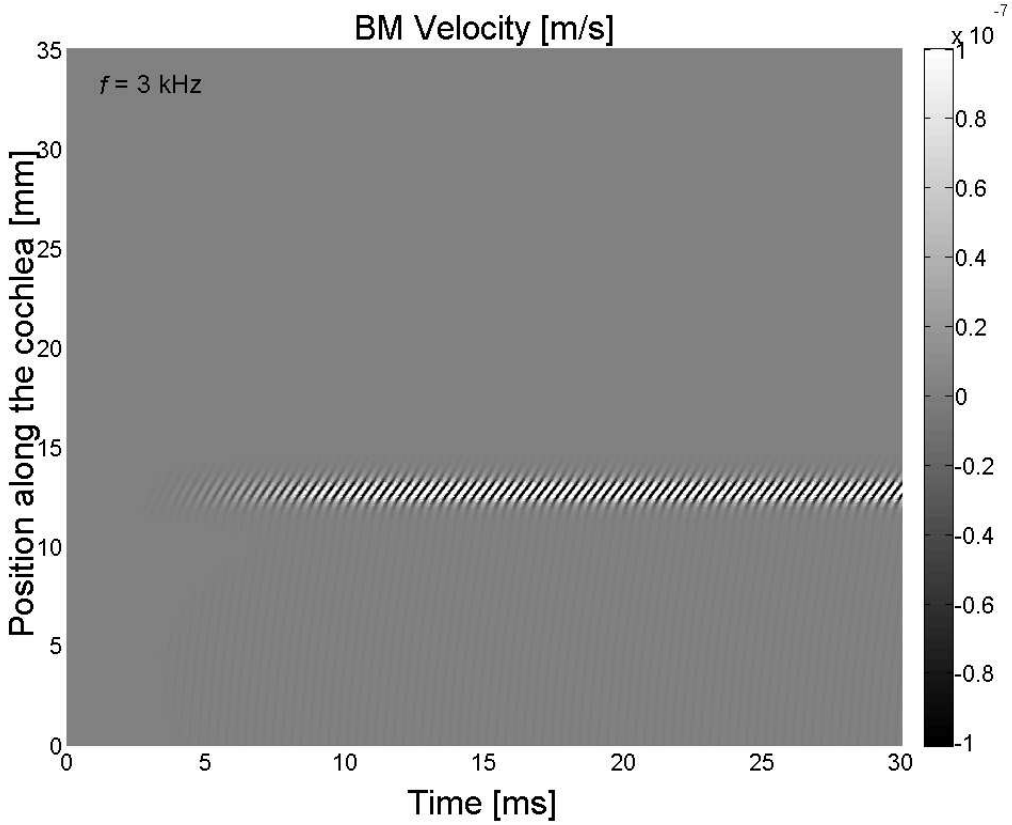


Figure 3.14: The BM velocity in a linear cochlea given a windowed 3 kHz stimulus tone at 0 dB SPL.

3.5.2 Nonlinear responses

The linear model simulates the cochlea's response to low-level stimuli where the CA is working at full strength. However, when the driving amplitude is increased, the OHC feedback force in the biological cochlea begins to saturate as it cannot accommodate the increased output requirements. Thus, the active contribution of the CA decreases relative to the driving level; this effect is sometimes referred to as 'self-suppression' in the literature (e.g. Kanis and de Boer, 1994). In order to model the nonlinear saturation of the feedback force, a Boltzmann function is applied to the displacement input of the feedback loop that determines the OHC force in the time domain. In effect, this acts as an automatic gain controller.

There are several reasons why the Boltzmann function was chosen for this task. Its shape well-approximates the input-output (stimulus pressure to intracellular voltage) characteristics of OHCs measured in isolation, *ex vivo* (Cody and Russell, 1987; Kros *et al*., 1992). However, the exact characteristics of the OHC response *in vivo* are as yet

unknown; hence, some degree of laxity in this aspect of the model can be allowed. The key features of this function are that it saturates with increasing input amplitudes, and that the saturation is non-symmetric. The latter characteristic is required to account for the presence of even-order components that have been measured in the harmonic distortion of the BM response (Cooper, 1998).

The first order Boltzmann function is given as

$$f(u) = \frac{\delta}{1+\beta e^{-u/\eta}} - \frac{\delta}{1+\beta}, \tag{3.63}$$

where u is the input displacement in units of distance; δ sets the saturation point, also in units of distance; β is a dimensionless quantity that scales the asymmetrical nature of the function; and η affects the slope of the function. In order to linearise the function for small input displacements, u, it is necessary to constrain

$$\eta = \frac{\delta\beta}{(1+\beta)^2}, \tag{3.64}$$

to yield a slope of 1 at small input displacement values, u. This was determined by equating the derivative of the Boltzmann function with respect to the input, u, to unity while evaluating at $u = 0$ and solving for η. Figure 3.15 illustrates the Boltzmann function and its slope, given parameter values of $\delta = 1$ and $\beta = 3$.

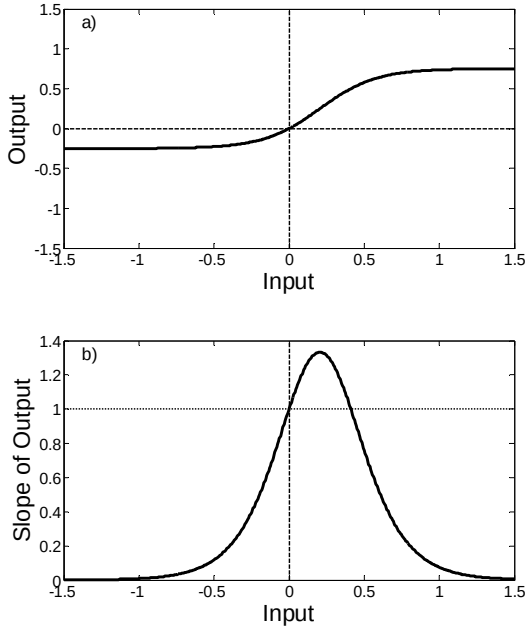


Figure 3.15.a-b: Boltzmann function characteristics: a) output vs. input, b) slope of output vs. input. Free parameter values are set as follows: $\delta = 1$ and $\beta = 3$.

While the saturation point of the Boltzmann function, δ, is set to 1 in Figure 3.15 for illustration purposes, it is varied as a function of position in the nonlinear cochlea. As different positions along the cochlea are driven with a constant pressure, each micromechanical element will displace a different amount. It is the relative displacement between the BM and the TM that is passed through to the saturation function in the nonlinear time domain simulation, and thus $\delta(x)$ has considerable bearing upon the results of the simulation. For example, if $\delta(x)$ is set to a constant value, the response at the base will be much less saturated than locations more apical; this is because the basal region is much stiffer and thus the micromechanical elements displace less for a given stimulus.

In order to generate a sensible distribution of $\delta(x)$, the maximum displacement at a given location across frequency in the linear model was used as a template for locations approximately 6 mm $\leq x \leq$ 27 mm; below 6 mm and past 27 mm, the value of $\delta(x)$ is fixed. This distribution is normalised to the maximum value, and scaled by 1 nm. The final distribution of $\delta(x)$ is shown in Figure 3.16.

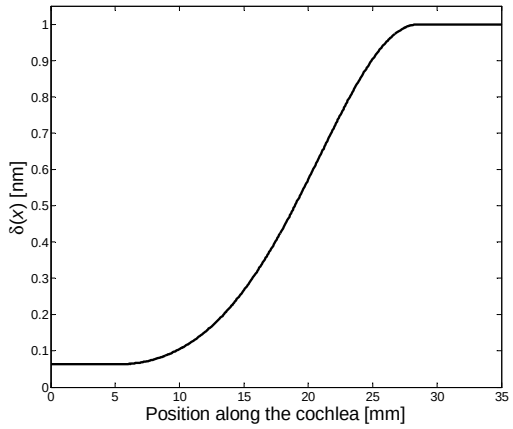


Figure 3.16: Nonlinear saturation point as a function of position, $\delta(x)$.

As with the linear simulations of pure tones, the first 10 ms of the stimulus is scaled by a half-Hanning windowed. Again a 3 kHz tone is applied, but now at varying intensities, varying from 0 dB SPL to 90 dB SPL in 3 dB steps. Figure 3.17 shows the mesh of the first 30 ms of BM velocity given a 45 dB SPL tone and a 90 dB SPL tone. The most striking difference between Figure 3.17.a and the linear simulation of Figure 3.14 is the appearance of a second region where the TW is peaking and travelling slower than adjacent regions centred at approximately 7 mm.

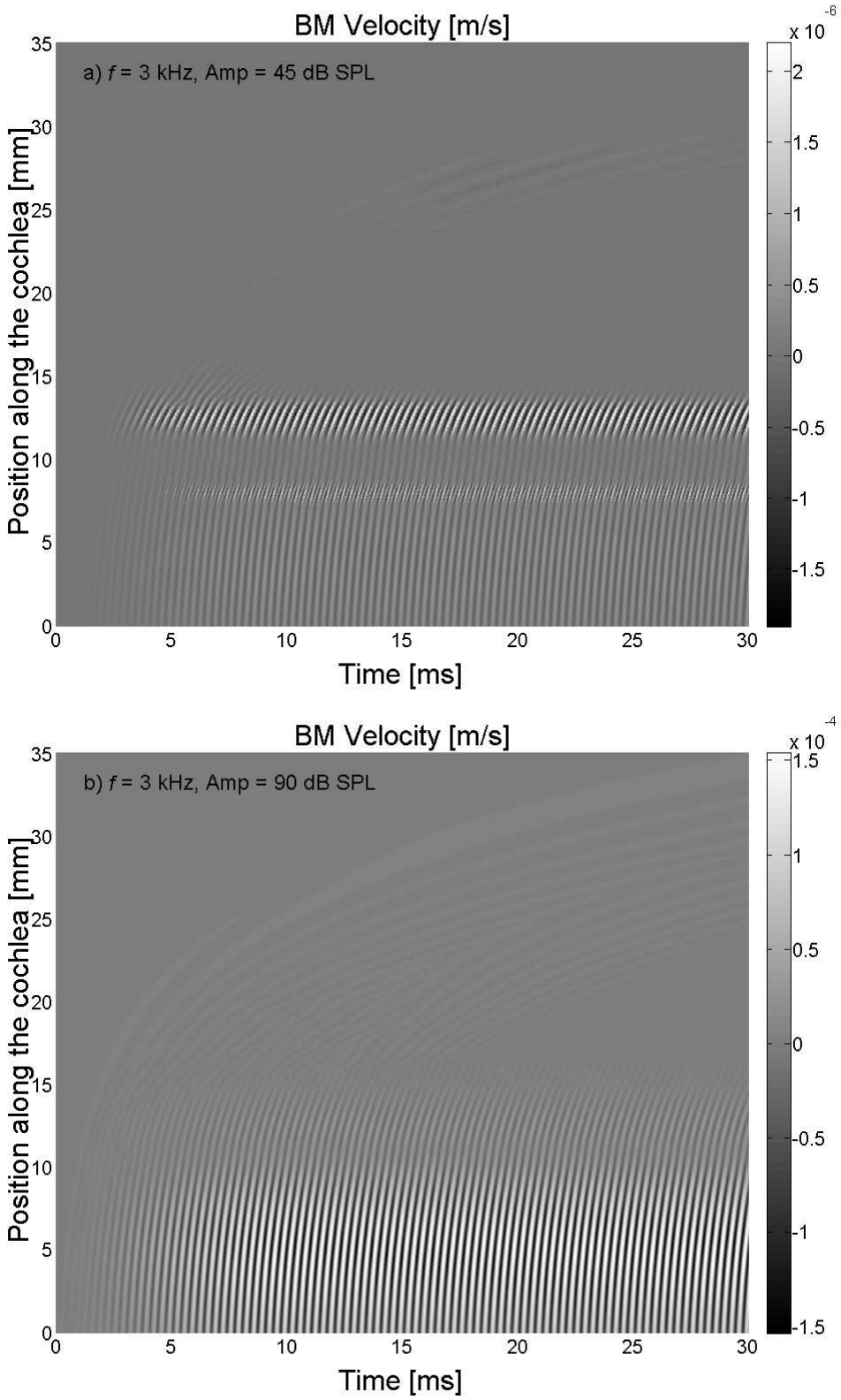


Figure 3.17.a-b: Mesh of the first 30 ms of BM velocity given a 3 kHz input tone at a) 45 dB SPL and b) 90 dB SPL in a nonlinear cochlea.

In contrast, Figure 3.17.b shows a much wider spatial excitation pattern at a stimulus level of 90 dB SPL, though a smaller peak is still visible near the 3 kHz place. Figure 3.18 plots the modulus of the overall BM velocity responses at various positions along the cochlea for excitation levels that vary from -20 to 100 dB SPL on a single graph. These calculations are made over the last 30 ms in a 100 ms simulation. Each curve represents the maximum steady state BM velocity as a function of position.

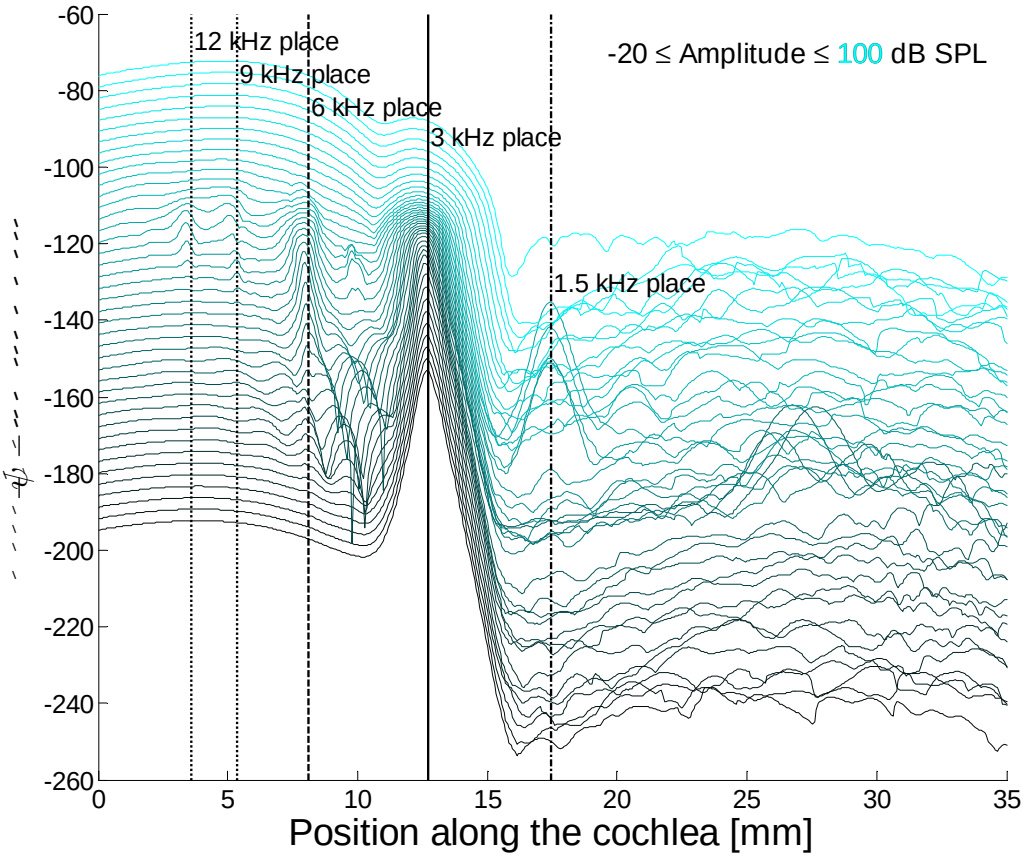


Figure 3.18: Growth of the steady state BM response given increasing amplitudes at 3kHz in 3 dB steps. Dashed vertical lines indicate the best places at 1.5, 3.0, 6.0, 9.0 and 12.0 kHz.

By visual inspection, Figure 3.18 shows that the response basal to the 3 kHz place grows linearly for much of the range, while the peak of the response grows much more compressively; this is consistent with the measured responses as described by Johnstone *et al*. (1986), for example. There are also numerous peaks that occur at seemingly-regular

intervals along the cochlea. Vertical lines at the best places of multiples of the fundamental tone indicate that these peaks are likely due to distortion.

Figure 3.19.a-d shows the normalised growth of several different harmonic components of the BM response, as obtained by taking the Fourier Series of the BM velocity at each location.

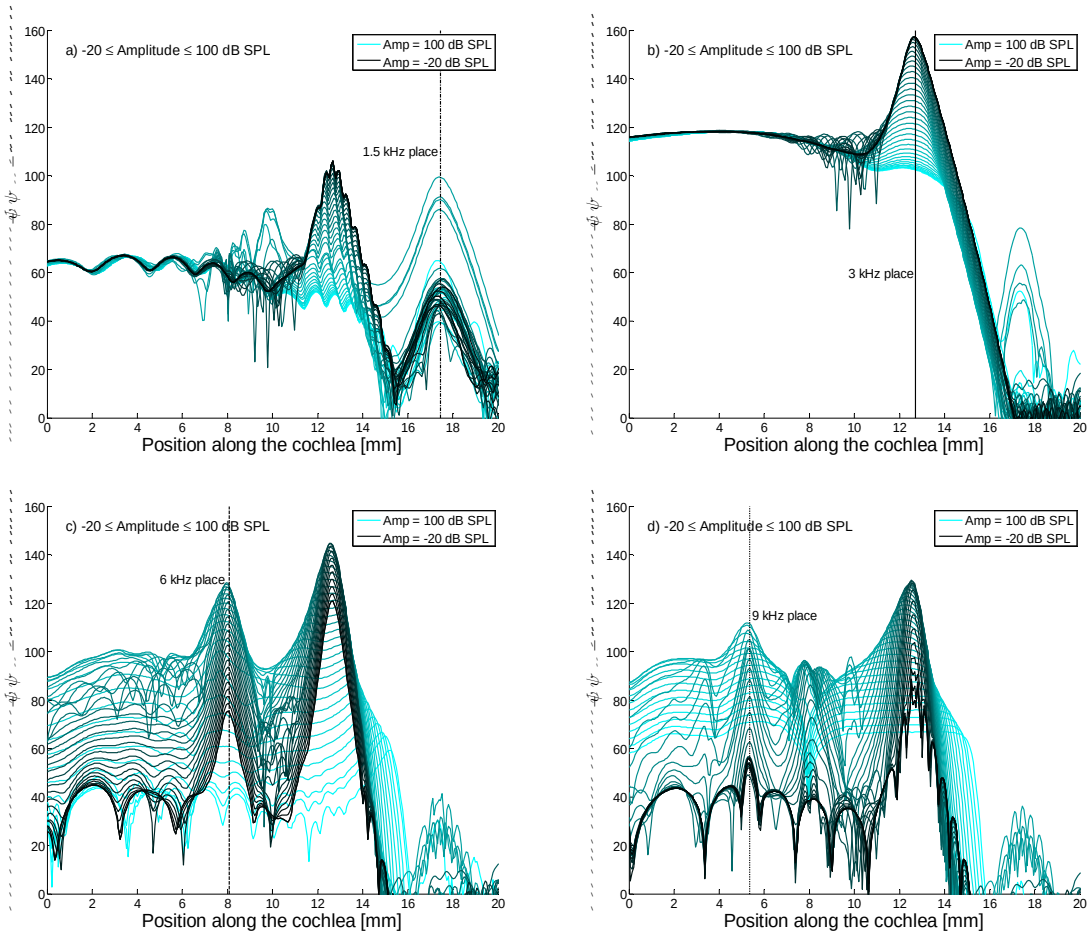


Figure 3.19.a-d: Normalised growth curve of a) the 1.5 kHz component, b) the 3 kHz component, c) the 6 kHz component, and d) the 9 kHz component of the BM response as a function of position along the cochlea.

It is important to remember that only a 3 kHz tone is being applied; the various components at 1.5, 6 and 9 components shown in Figure 3.19 seem to be a result of harmonic distortion. The various harmonic components of the BM response also grow at different rates. Figure 3.20 plots the growth of these components at several positions along the cochlea.

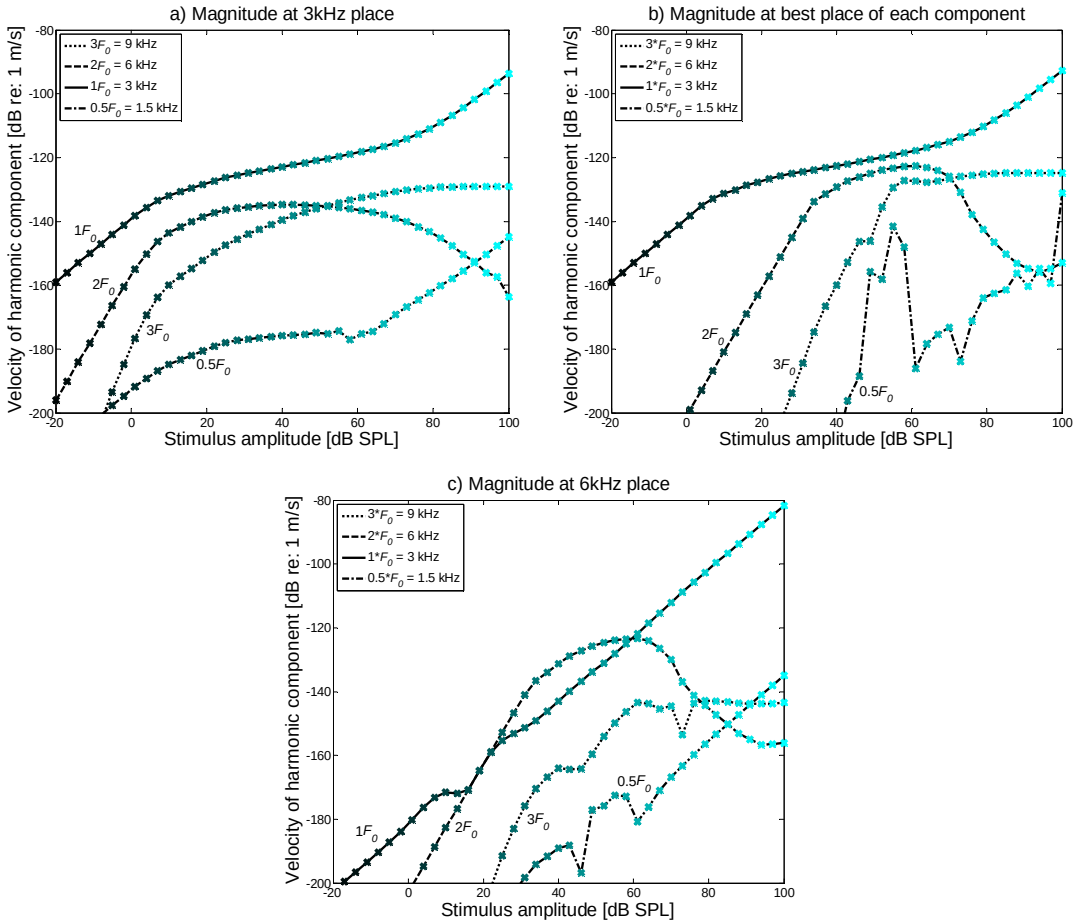


Figure 3.20.a-c: Growth of distortion components generated a) at the 3 kHz place; b) near (±~1 mm about) the best places of the distortion frequencies; c) at the 6 kHz place.

Figure 3.20 shows that the nonlinear cochlear model has compressed 120 dB of stimulus intensities at the 3 kHz harmonic component into a much smaller dynamic range, approximately 60 dB of BM motion. The 3 kHz level curve appears to have corner amplitudes of approximately 5 dB SPL and 70 dB SPL where the BM response transitions from linear growth to compressive growth, and then back to linear growth. In order to match the growth curve to commonly measured values in animals where corner amplitudes are approximately 30 and 90 dB SPL, as given by Pickles (2003) for instance, the saturation point along the entire cochlear model could be shifted up by a factor of 10. However, it would be advisable to complete a more exhaustive study of growth curves at different frequencies (and thus locations along the cochlea) before adjusting simulation parameters as this is very recent work. It may be that the growth curve of the model's

fundamental response at more basal regions of the cochlea, where the experimental data were obtained, matches the measured data.

Figure 3.20.a is comparable to Figure 4 of Cooper's (1998) paper regarding harmonic distortion, where BM distortion components were measured in a live guinea-pig cochlea. As with the data presented in that article, the growth of the distortion products is strongly compressive. The $2f_0$ component of Figure 3.20.a shows a marked decrease in its absolute magnitude as the stimulus level is increased which is not observed in Cooper (1998). However, the decrease in the $2f_0$ component appears to start approximately at the second corner amplitude of the f_0 component, where the response becomes linear again; this high-level region was not measured in Cooper (1998).

In Figure 3.20.b, the $2f_0$ and $3f_0$ components grow at a rate of approximately 2 dB/dB at the best places of these frequencies, before beginning to show signs of compression at approximately 30 and 50 dB SPL, respectively. Figure 3.20.c is directly comparable to Figure 3A of Cooper (1998), which displays harmonic data of BM motion in a guinea pig at a location whose best frequency is twice the stimulus frequency ($f_{cf} = 0.5f_{tone}$). The growth of the 3 kHz driving tone component is nearly linear in Figure 3.20.c and the 8 kHz driving tone is completely linear in 3A of Cooper (1998). The growth of the $3f_0$ component is qualitatively equivalent in both plots as well. However, the magnitude of the $2f_0$ component in Figure 3.20.c exceeds that of the fundamental between 25 and 60 dB SPL, whereas it is always less than the fundamental in 3A of Cooper (1998). The latter effect may be due to the choice of the β parameter here, which shapes the vertical offset of the Boltzmann function, though further work is necessary in this area.

Based upon the simulations above, it seems plausible that Cooper was correct in proposing that harmonic distortion is primarily *distorted amplification,* and not *amplified distortion*. However, it is also possible that distortion generated at one site may propagate to its best place and become amplified there. This is even more likely in a cochlear model that contains inhomogeneities in the CP impedance as a function of position, as these perturbations could result in the reflection of energy. Kemp's (1978) nonlinear wave-fixed theory also assumes that the TW itself modifies the cochlear mechanics in an amplitude-dependent manner such that backward-travelling reflections could arise from this local irregularity in the CP impedance.

The simulation of nonlinear cochlear responses in the time domain is an area that requires a considerable amount of further work. As discussed at the start of this chapter, the most significant downside of such investigations is the computational cost. The total cpu time required for the thirty-one simulations of 100 ms of data was approximately 490 hours. The total duration of a given simulation is a complicated function of applied error tolerances and the relative magnitude of the motion in the cochlea, the latter of which is dependent on driving amplitude and stability. In order to calculate the model's response at a single position to a wide range of frequencies and stimulus levels, as is sometimes performed experimentally (see Figure 1 of Cooper (1998), for example), many hundreds of computer-hours would be the required.

In the next chapter, instabilities in the cochlear model are studied in terms of their generation mechanisms and characteristics. Further time domain simulations of unstable cochleae continue to show that there is much greater complexity in the response of nonlinear models than is predicted by linear theory.

Chapter 4

Spontaneous Otoacoustic Emissions

Calculating the stability of Neely and Kim's (1986) frequency domain model of the cochlea was the primary motivation for reformulating it in state space. However, with the state space model came the ability to test not only if the linear system was stable or not, but to actually study the conditions that give rise to the instabilities. The resultant frequencies and characteristics of the instabilities also give insight into the system.

One of the initial tests of stability included perturbing the feedback gain at various locations with a step-discontinuity. The instabilities that emerged from these investigations appeared in an arc of poles, where the most unstable frequency was close to the characteristic frequency (CF) at the location of the discontinuity. The frequencies of the poles in the arc seemed to fall at regular intervals; this was of particular interest, as it has been reported that spontaneous otoacoustic emissions (SOAEs) in a single individual exist at regular intervals in mammals (Dallmayr, 1985, 1986; Talmadge *et al.*, 1993; Braun, 1997). Thus, in order to compare the model's response with experimental results in humans, it was decided that the parameters for Neely and Kim's (1986) model of a cat cochlea would be revised to account for the characteristics of a human cochlea.

Two sets of revised model parameters have been used to study the distribution of linear instabilities. The first replicated the most commonly observed spacing between SOAE frequencies in humans, and seemed to support one of the prevailing theories regarding SOAE generation; these results were published in Ku *et al.* (2008), which can be

found in Appendix C.2. However, in order to match the variation of spacings between adjacent human SOAEs as a function of frequency, as described by Shera (2003), a second set of parameters was devised. The latter set of parameters has formed the basis of the results presented in this thesis.

This chapter begins by describing the findings presented in Ku *et al.* (2008). This work was primarily an investigation into the generation mechanisms of instabilities in the cochlea. Subsection 4.1 introduces various theories that have been proposed to explain the formation and characteristics of SOAEs. Subsection 4.2 shows how the stability of the model is affected given various perturbations in the micromechanical feedback gain as a function of position. Some details of the results here differ from those presented in Ku *et al.* (2008) given the application of the newer set of parameters, though the underlying conclusions are still valid. Findings from more recent investigations are presented in subsection 4.3 in the form of nonlinear time domain simulations of unstable cochleae. The results of these newer simulations are discussed in the context of the linear theory of SOAE generation.

4.1 Theories of SOAE generation

SOAEs are believed to be a feature of a normally functioning CA, as they are commonly detected in an estimated range of 33% to 70% of all normally-hearing ears (Talmadge *et al.*, 1993). Where SOAEs are detected, stimulus frequency-, distortion product- and transient evoked- otoacoustic emissions (SFOAEs, DPOAEs and TEOAEs) are also often present. There is evidence to suggest that all forms of OAEs are related and directly tied to the sensitivity of hearing (Zwicker and Schloth, 1984; McFadden and Mishra, 1993; Talmadge and Tubis, 1998; Shera and Guinan, 1999; Kalluri and Shera, 2007). Two primary classes of cochlea-based theories regarding the production of SOAEs are discussed below: a local-oscillator model, and a distributed backscattering concept.

Gold (1948) first formed the basis of a local-oscillator model of SOAE generation when he proposed that a perturbation may ‘bring an [active] element into the region of self-oscillation, when it is normally so close to [instability].’ Evidence in the literature suggests that SOAEs are associated with BM oscillations. For example, Nuttall *et al.* (2004) measured a SOAE that had a counterpart in spontaneous mechanical vibration of the BM at the same frequency in a guinea pig. Further work performed by Martin and

Hudspeth (2001) considered how locally unstable elements of the CA may be responsible for SOAEs. However, without careful tuning, a local-oscillator model fails to account for the regular spacings between unstable frequencies observed in mammalian SOAEs.

The strong peak in the distribution of log-normalised spacings between adjacent SOAE frequencies, termed the preferred minimum distance (PMD), has been demonstrated by various studies (Dallmayr, 1985, 1986; Talmadge *et al.*, 1993; Braun, 1997). A similar value is found in the average frequency spacings between the spectral peaks of SFOAEs and TEOAEs when measured in the ear canal (Zwicker and Schloth, 1984; Shera, 2003). The PMD corresponds to a frequency spacing of approximately 0.4 Bark, or a distance of about 0.4 mm along the human cochlea (Dallmayr, 1985, 1986). Most SOAEs occur in the range of 0.5-6 kHz (Probst *et al.*, 1991) and demonstrate the PMD, though Zweig and Shera (1995) and Shera (2003) showed that the average spacings of both SOAEs and the spectral peaks of SFOAEs measured in the ear canal do vary somewhat with frequency.

Strube (1989) argued that a periodic variation or 'corrugation' in the micromechanical parameters was responsible for the observed PMD in SF- and TE-OAE measurements in the ear canal. This was said to be the result of distributed backscattering of the travelling wave (TW) similar to the phenomenon of Bragg reflection in a crystal. In this theory, the period of the corrugation must correspond to one-half of the wavelength of the TW, thus generating constructive interference at particular frequencies. Kemp (1979) also proposed a theory of SOAE generation which assumed a distributed backscattering mechanism; his theory required multiple internal reflections of forward- and backward-travelling waves between the middle ear boundary and an inhomogeneous region of the cochlea.

Since Kemp (1979) first presented the idea, numerous authors have made contributions to the multiple-reflection theory (Zwicker and Peisl, 1990; Zweig, 1991; Shera and Zweig, 1993; Talmadge and Tubis, 1993; Zweig and Shera, 1995; Talmadge and Tubis, 1998; Shera and Guinan, 1999; Shera, 2003). Shera and Zweig (1993) proposed that a spatially dense and random array of reflection sites exists along the entire cochlea which acts in concert with the middle ear boundary to form standing waves, which Shera (2003) likens to a laser cavity. This concept was fully developed in Zweig and Shera (1995). Though energy is reflected at all frequencies by a perturbation along the cochlea,

wavelets scattered from forward-travelling waves that peak in the region of the inhomogeneity dominate the response, since the amplitude is highest there.

For an active standing wave resonance to develop in this multiple-reflection theory, the spatial distribution of inhomogeneities in the given region must contain components at the wavenumber that creates constructive interference with the incoming wave, just as with Bragg scattering (Shera and Zweig, 1993; Zweig and Shera, 1995). Further requirements include an active region between the middle ear boundary and the reflection site to overcome the viscous damping in the cochlea, and a TW frequency that undergoes an integer number of cycles of round-trip phase change between the middle ear and the cochlear reflection site; this naturally gives rise to the PMD in SOAEs measured in the ear canal. However, the existence of a spontaneous oscillation in the cochlea does not guarantee its detection as an SOAE; it must also remain sufficiently powerful to be measurable in the ear canal after transmission through the middle ear.

An alternative theory suggests that irregular middle ear transmission characteristics may be a cause of some OAEs (Nobili *et al.*, 2003). However, the numerical accuracy of these simulation results has been contested elsewhere (Shera *et al.*, 2003), and such irregularities are not often reported. For the purposes of this investigation, a smooth middle ear boundary is implemented and only cochlea-based theories of SOAE generation are discussed.

The goal of subsection 4.2 is to test whether the predictions of Zweig and Shera's (1995) multiple-reflection theory of SOAE generation are observed in a mathematical model of linear cochlear mechanics. Previous work has relied upon phenomenological methods (Zweig and Shera, 1995; Shera, 2003), or multiple time domain simulations (Talmadge *et al.*, 1998), to support this theory. In contrast, a state space formulation of the cochlea (Elliott *et al.*, 2007) is used here that is capable of rapidly and unambiguously calculating the unstable frequencies in a given linear model. This method is thus especially well-suited to generating the large number of results from individual cochleae necessary to ensure statistically significant data.

It should be noted that only the linear stability of the cochlear model is considered in subsection 4.2. In a biological cochlea, the amplitude of an instability would eventually stabilise due to the natural saturation of the feedback force generated by the CA.

Furthermore, it is possible that the number of SOAEs predicted by the linear model could change in a nonlinear model due to distortion or suppression; this is tested in Section 4.3.

4.1.1 Theoretical predictions

According to Shera and Zweig (1993), the average distance between resonant positions of SOAEs along the cochlea is

$$\overline{\Delta x}_{SOAE} \approx \frac{1}{2}\lambda_{peak}, \tag{4.1}$$

where λ_{peak} is the wavelength of the TW in its peak region. This is a departure from previous predictions in that the assumed distance between SOAE characteristic places is determined by the wavelength of the TW and not a specific corrugation in the mechanical parameters. The wavelength of the model's TW as a function of position and frequency is given by equation (2.35); thus, the wavelength at the peak for a given location is determined by substituting in the CF, as shown by (2.36).

Given that the frequency to place map in the cochleae of mammals is roughly exponential, the PMD between SOAE frequencies is

$$f/\Delta f \approx 2l/\lambda_{peak}, \tag{4.2}$$

where l is the cochlear length scale, the distance over which the best frequency changes by a factor of e. It is also possible to define a log-normalised spacing between two adjacent SOAE frequencies, f_a and f_b, as the ratio of their geometric mean divided by their difference,

$$f/\Delta f = \frac{\sqrt{f_a f_b}}{|f_a - f_b|}. \tag{4.3}$$

The PMD in humans is approximately 15 when expressed in terms of $f/\Delta f$, though this value increases somewhat with SOAE frequency (Shera, 2003). Figure 4.1 shows the length scale, l, and the wavelength of the TW at its peak, λ_{peak}, for the model used in this thesis. Together, these quantities yield the predicted $f/\Delta f$. The length scale generated by this set of parameters is roughly the same as that generated by the parameters in Ku *et al*. (2008). However, λ_{peak} is an increasing function of position along the cochlea in the new model, whereas $\lambda_{\text{peak}} \approx 0.9$ mm for all but the most apical regions of the cochlea in the model of Ku *et al*. (2008).

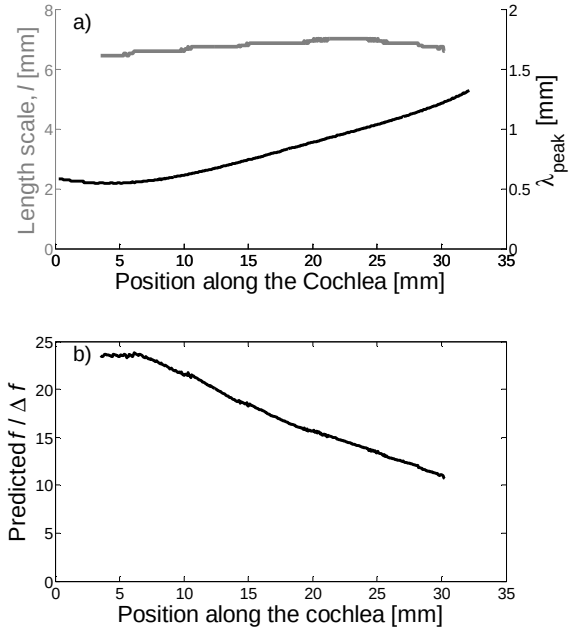


Figure 4.1.a-b: The length scale and wavelength of the TW at its peak, shown in panel (a), combine to yield the predicted $f/\Delta f$, shown in panel (b).

4.2 Linear stability given perturbations in γ(x)

It has been previously reported that deviations from a smoothly varying set of micromechanical parameters can cause instability in cochlear models. It is believed that the frequencies of cochlear instability represent the frequencies of potential SOAEs. Elliott *et al.* (2007) demonstrated that these models are most sensitive to abrupt changes in the gain as a function of position. In the current investigation, greater consideration is given to the nature of the inhomogeneities introduced and the resultant characteristics of the unstable frequencies. The feedback gain as a function of position along the cochlea, $\gamma(x)$, has been chosen as the parameter to be perturbed. In order to compare the relative level of instability present in a cochlea, it is instructive to examine the number of unstable frequencies present. However, to further quantify the magnitude of a cochlear model's instability, the concept of a pole's damping ratio is applied, as reviewed in subsection 3.2.1.

The damping ratio of an unstable pole is useful as it relates the frequency-independent rate at which the system will become unstable; the average value of many poles can also be compared across different cochlear models. This quantity is referred to

as the *undamping ratio* in this paper, in the context of discussing unstable poles, and is assigned the lowercase Greek letter final sigma:

$$\varsigma = -\zeta \tag{4.4}$$

A step change in gain is employed as a starting point for the discussion of cochlear stability analysis. From there, sinusoidal spatial variations and the band-limited random spatial variations are applied as gain distributions. It is important to note that the step- and sinusoidal- distributions of $\gamma(x)$ are introduced to understand the underlying mechanisms of SOAE generation and should not be interpreted as an attempt to model what exists in a healthy human cochlea.

4.2.1 Step changes in γ(*x*)

A step change in gain gives rise to a discontinuity in the smooth variation of BM impedance as a function of position along the cochlea. An ideal step in space has a well-distributed wavenumber spectrum, and thus should reflect wavelets across a wide range of wavelengths. One additional consequence of varying the gain as a function of position, $\gamma(x)$, is that the underlying properties of the TW are affected. For instance, a higher gain results in a shorter λ_{peak}. To minimize this effect, a relatively small amplitude step was chosen with a ±2.5% deviation from nominal gain on either side of the step. The stability plot for the cochlear model with such a step imposed on the gain at 13.6 mm from the base of the cochlea is shown in Figure 4.2.

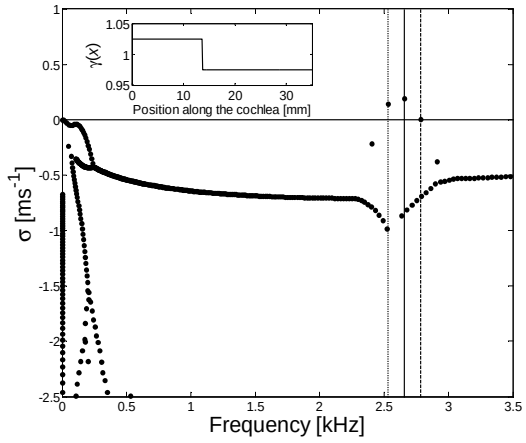


Figure 4.2: Stability plot for a cochlea with the stepped gain as a function of position inset: $\gamma(x < 13.6 \text{ mm}) = 1.025$ and $\gamma(x \geq 13.6 \text{ mm}) = 0.975$. Note the frequency scale has been shortened to clarify the locations of the unstable poles. Vertical lines indicate the frequencies of the instabilities: at 2.533 (dotted), 2.659 (solid) and 2.784 kHz (dashed).

Three distinct frequencies are found to be unstable in this cochlea, at 2.533, 2.659 and 2.784 kHz. These frequencies are all close to the CF at the location of the discontinuity, which is 2.640 kHz. According to Zweig and Shera (1995), only the frequencies whose responses peak in this region may become unstable since not enough energy is reflected otherwise; this is seen in Figure 4.2 as only three frequencies near the CF at the position of the discontinuity are unstable. Furthermore, there is a range of successively more stable poles that follow an arc leading away from the three unstable poles, both higher and lower in frequency. Presumably, the TWs of these frequencies are not reflected strongly enough by the discontinuity to cause instability.

The resultant spacings between the two pairs of adjacent unstable frequencies, $f/\Delta f$, are approximately 20.5 for the pair lower in frequency, and approximately 21.7 for the pair higher in frequency. This is consistent with the expectations given a slightly lower γ value apical of the discontinuity, and a slightly higher γ value basal to the discontinuity. To better understand why only these specific frequencies become unstable, Figure 4.3 shows the magnitudes and phases of the BM velocity responses at these frequencies when a nominal value of gain is used throughout the cochlea, $\gamma(x)=1$.

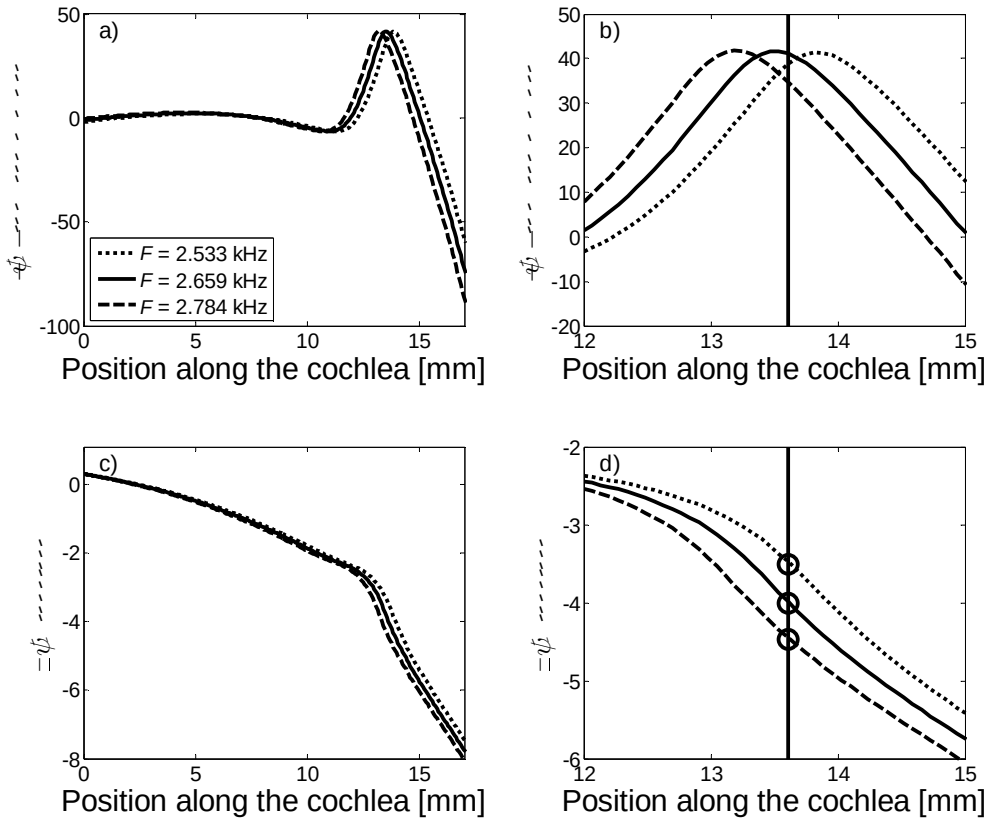


Figure 4.3.a-d: Magnitude (a, b) and phase (c, d) of BM velocity for excitation at 2.533 (dotted), 2.659 (solid) and 2.784 kHz (dashed) given a baseline model with nominal gain, $\gamma(x) = 1$. Panels (b) and (d) show expanded axes for clarity of interpretation. A vertical line is drawn at the location of the discontinuity of Figure 4.2 in the zoomed-in panels. Circles in the phase plot (d) indicate phase shifts of -3.5, -4.0 and -4.5 cycles at this location.

A vertical line through Figure 4.3.b and Figure 4.3.d denotes the location along the cochlea of the discontinuity applied in Figure 4.2. This line intersects with the phase responses of the 2.533, 2.659, and 2.784 kHz stimulus tones at -3.5, -4.0, and -4.5 cycles, respectively, within an accuracy of 1%. This is consistent with the 'cochlear laser' theory of SOAE generation which states that the phases of the unstable frequencies must undergo an integer number of cycles of total phase change between the reflection site and the middle ear boundary in order to combine constructively over successive reflections. For the unstable frequencies shown above, the 'round-trip' phase change would equal 7, 8, and 9 cycles. Re-examining Figure 4.2 in light of this feature, the stable poles that follow the same arc as the unstable poles must also represent frequencies that scatter wavelets which

constructively combine, but perhaps are too weak to overcome the damping basal to the inhomogeneity.

The total number of poles of a cochlear model is dependent upon the number of micromechanical elements; each additional element introduces four more poles. However, the multiple-reflection theory of SOAE generation predicts that the number of instabilities created by a reflection site is dominated by the rotation of the TW phase and the magnitude of the reflected wave. As such, only a finite number of instabilities should appear near the CF of the reflection site, regardless of the number of elements used in the model. This is tested in Figure 4.4 by applying the same spatial step discontinuity in $\gamma(x)$ to several cochlear models with different numbers of micromechanical elements, N.

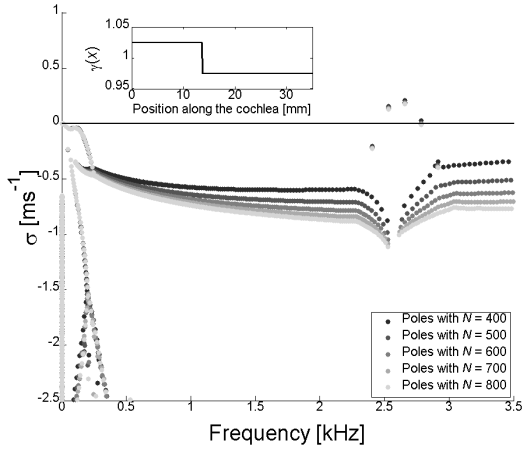


Figure 4.4: Stability plot for several cochlear models with varying values of N, the number of micromechanical elements. A stepped gain as a function of position was applied to each model and inset in the figure: $\gamma(x < 13.6 \text{ mm}) = 1.025$ and $\gamma(x \geq 13.6 \text{ mm}) = 0.975$. Note the frequency scale has been shortened to clarify the locations of the unstable poles.

Figure 4.4 shows the variation of the stability plot as N is increased from 400 to 800. The same step discontinuity is applied, which results in three distinctly unstable frequencies in each model, independent of N. Note however that the positions of all the poles do vary somewhat, tending toward greater stability with increasing N. The pole positions converge as N is increased. Unfortunately, simulating models of greater and greater numbers of elements increases the computational load dramatically. Furthermore, the fundamental stability results do not vary a great deal, as shown above; as such, a compromise of $N = 500$ appears suitable for the purposes of the current investigations.

4.2.2 Sinusoidal variations in γ(*x*)

A distribution of gain that is sinusoidal as a function of position is of interest because its wavenumber spectrum is concentrated close to a single wavenumber, just as a sinusoidal waveform that is a function of time has a frequency spectrum that is concentrated close to a single frequency. This set of simulations follows the theory outlined by Strube (1989), which assumes uniform corrugations in gain along the BM. A range of wavelengths were chosen for the sinusoidal variation of gain as a function of position along the cochlea, denoted λ_{sin}. λ_{sin} is varied from approximately 1 mm down to 0.14 mm, the latter being the spatial Nyquist limit of the model. A 14% peak-to-peak variation in amplitude about nominal gain generated instabilities over this spatial range.

Figure 4.5 shows a number of stability plots of models as λ_{sin} is varied. At the longest wavelengths, panels (a) and (b) for instance, the upper arm of poles periodically becomes unstable in frequency. This large-scale periodicity appears to correspond to the spatial corrugation of the sinusoidal $\gamma(x)$. For instance, the characteristic places of 14.5 kHz and 16.9 kHz (the frequencies of the 'bumps' in Figure 4.5.a with λ_{sin} = 1 mm) are almost exactly 1 mm apart on the model's frequency-to-place map. However, there is also a smaller-scale periodicity within the spacings of unstable frequencies; the three unstable poles between 14.3 kHz and 14.77 kHz in Figure 4.5.a, for instance. This is likely due to coherent reflection in these bands where the peaks of the corrugation act as individual reflection sites, similar to the step-discontinuity results.

As λ_{sin} decreases, these general trends become indistinguishable, but both the number of instabilities and their average undamping ratios begin to vary; this is summarised in Figure 4.6. As expected, given the theories of Strube (1989), Shera and Zweig (1993), and Zweig and Shera (1995), the strongest instability occurred when the wavelength of the sinusoid, λ_{sin}, was approximately half the average peak wavelength; this value occurs at 0.32 mm in the model. In order to more fully understand the results of Figure 4.6, the same results as summarised in Figure 4.6 are presented for the unstable poles within half-octave bands in Figure 4.7 and Figure 4.8.

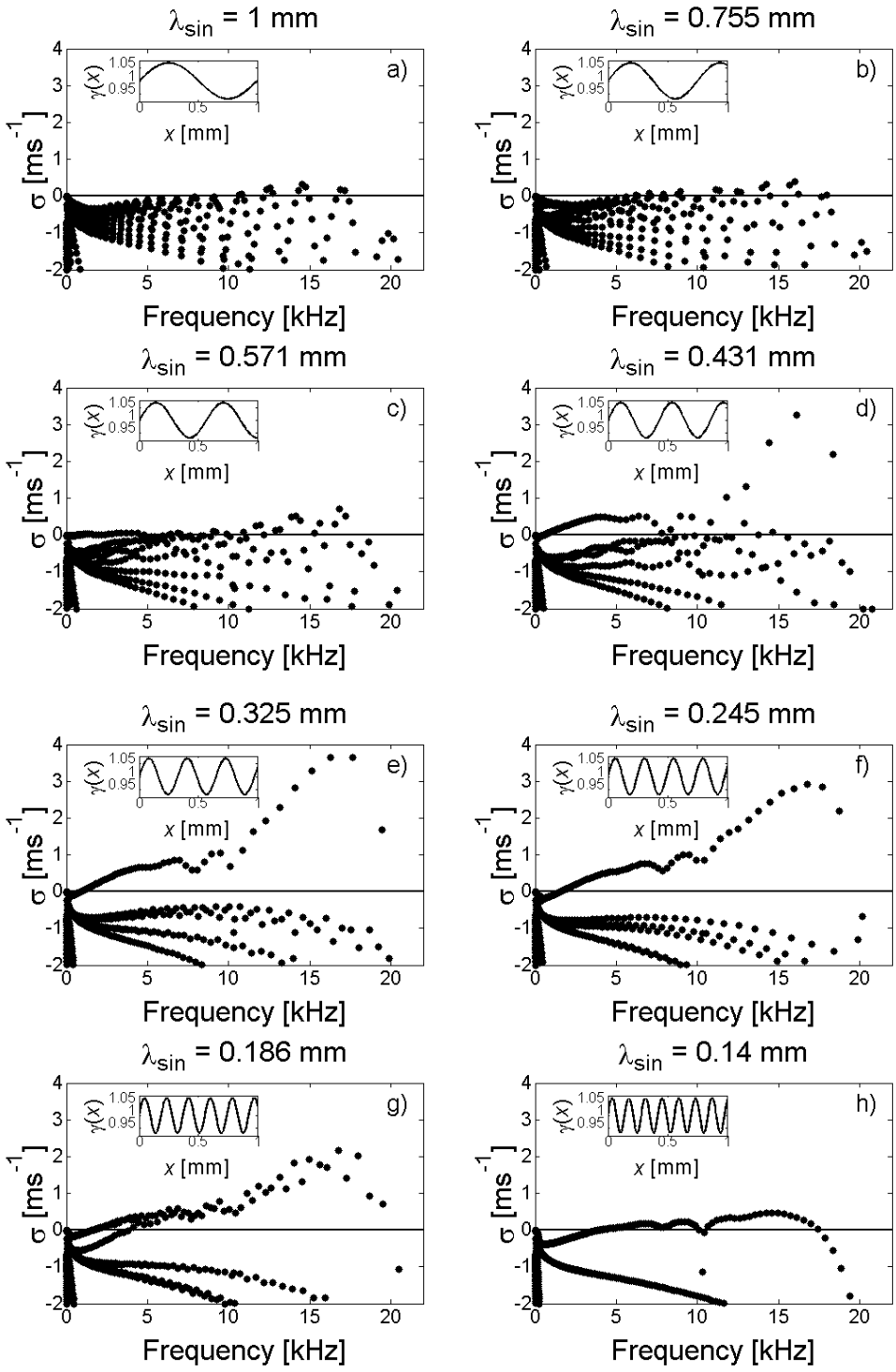


Figure 4.5.a-h: Stability plots models with 8 different values of λ_{sin}.

Note that the inset sinusoidal distributions of gain are sampled more densely than the discretisation size of the model for clarity.

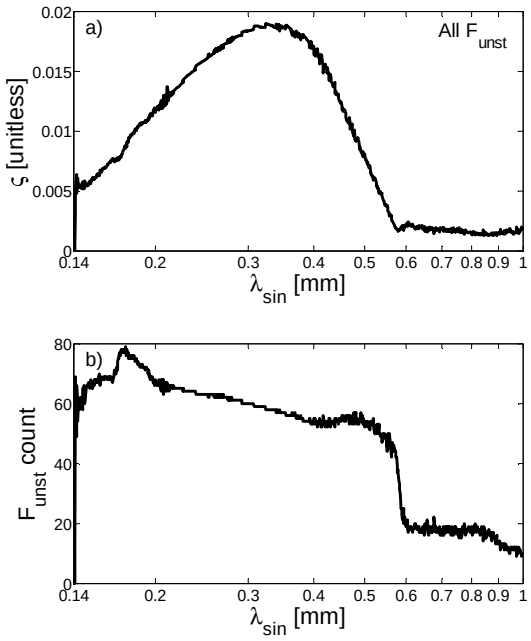


Figure 4.6.a-b: Average undamping ratio of the instabilities (a) and total number of unstable frequencies (b) given variations in the sinusoidal wavelength of $\gamma(x)$.

The results of Figure 4.7 and Figure 4.8 show that the sinusoidal wavelength that causes the most instability within a given band tends to decrease with increasing frequency. For instance, as shown in the banded undamping ratios of Figure 4.7, the 'characteristic' sinusoidal wavelength is 0.39 mm for frequencies below 2 kHz, and steadily decreases to 0.29 mm at the region of the 11.3 kHz band. The corresponding λ_{peak} within these frequency ranges is approximately 0.8 mm and 0.6 mm, respectively, roughly twice the characteristic sinusoidal wavelength; this is consistent with the multiple-reflection theory, as explained in (4.1). Summarized another way, the greatest instability is generated within a given band of frequencies when the reflection sites are exactly aligned to half the wavelength at its peak in that frequency range:

$$\lambda_{\sin} = \frac{1}{2}\lambda_{peak}(\omega) \Rightarrow \max\left(\varsigma(\omega)\right). \tag{4.5}$$

This trend breaks down somewhat in the highest frequency bands near the base of the cochlea where the model is closest to instability; this is also potentially due to spatial aliasing of λ_{sin} at small wavelengths.

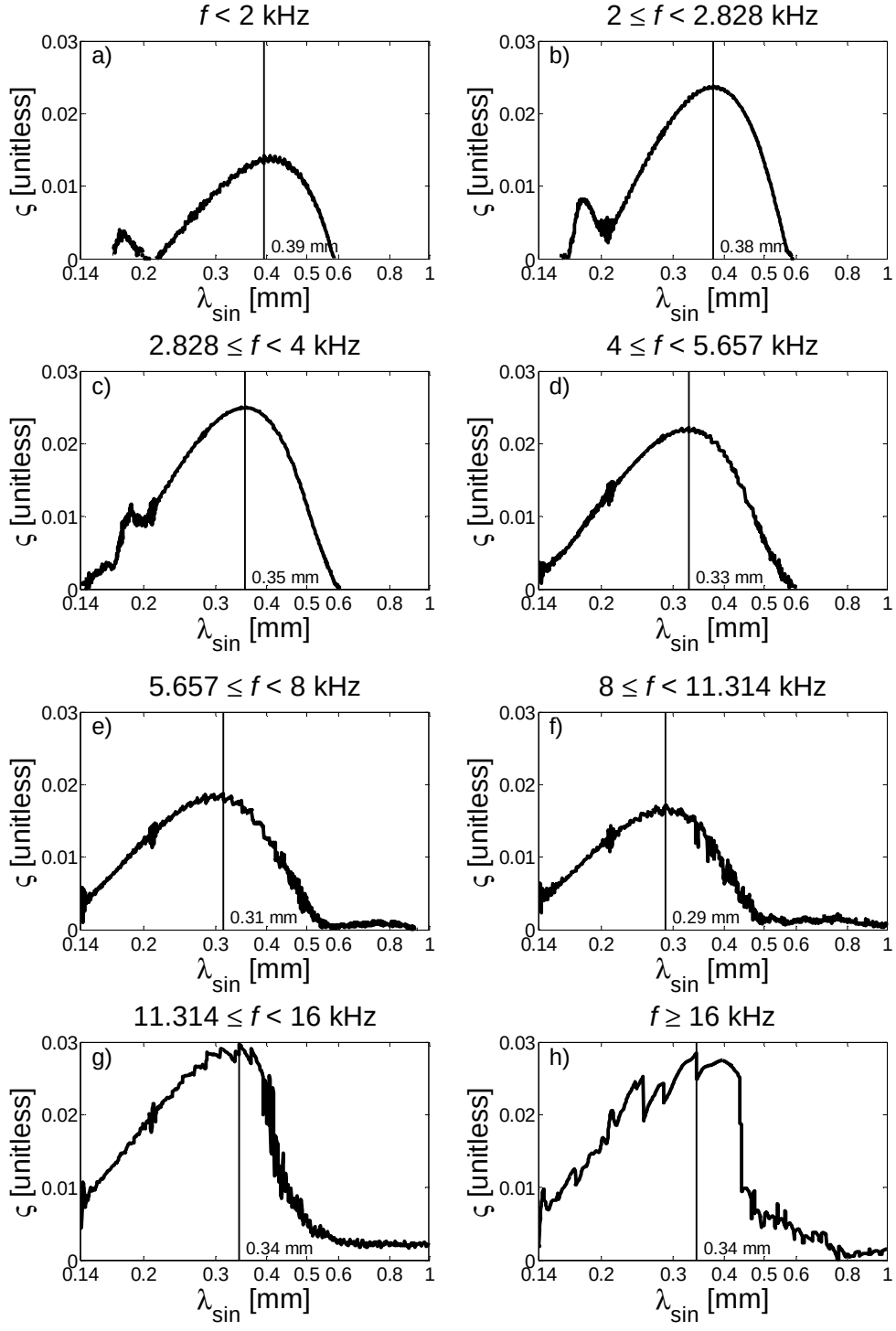


Figure 4.7.a-h: Average undamping ratio of the instabilities that fall within half-octave bands as λ_{sin} is varied. The sinusoidal wavelength that generates the strongest instability in each band is marked by a vertical line

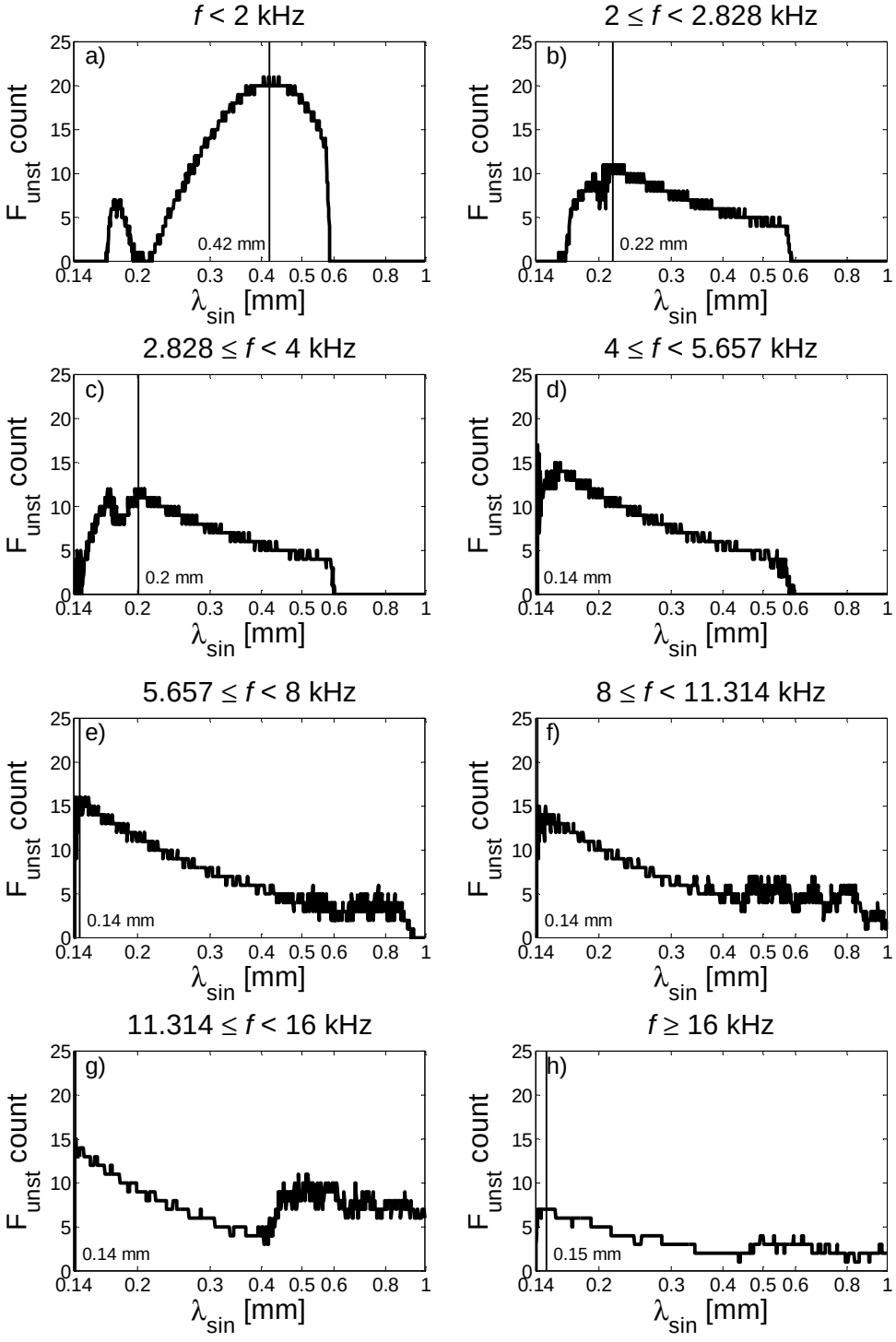


Figure 4.8.a-h: Total count of the instabilities that fall within half-octave bands as λ_{sin} is varied. The sinusoidal wavelength that generates the strongest instability in each band is marked by a vertical line.

Figure 4.8 shows that the sinusoidal wavelength at the peak in the number of instabilities also varies also with frequency, but these results are not as clean as Figure 4.8. This is possibly due to the magnitude of the sinusoid applied, as lower amplitudes may resolve more clearly the most sensitive portions of the cochlea. For example, the (a) and (b) panels show results toward the apex with the lowest levels of enhancement, and thus are only sensitive when $\lambda_{\sin} = 1/2\,\lambda_{peak}$. Furthermore, there is no instability for the frequency band where $f < 2$ kHz when the sinusoidal wavelength is approximately 0.2 mm. This sinusoidal wavelength, $\lambda_{\sin} = 1/4\,\lambda_{peak}$, may generate interference between adjacent reflection sites and thus decrease the model's tendency to instability. Again, this is not visible in the higher frequency bands, though this is possibly due to the spatial Nyquist limit of the model; sampling at higher spatial frequencies may resolve this trend more clearly.

4.2.3 Random variations in γ(*x*)

Shera and Zweig's (1993) theory of SOAE generation assumes that the cochleae of normal-hearing humans contain a dense but random array of inhomogeneities. Each of these place-fixed perturbations reflects energy from the forward TW (Talmadge *et al.*, 1993; Shera and Zweig, 1993; Zweig and Shera, 1995). In this section, the stability of cochlear models with band-limited, spatially random gain distributions is used to approximate what is postulated to exist in a human cochlea. A 5th order Butterworth filter was employed to band-limit gain distributions in the wavenumber domain (Lineton, 2001). The low-wavenumber cut-off frequency is fixed at the length of the cochlea itself, in order to prevent any DC shifts in the gain. The high-wavenumber cut-off frequency is initially set to 6.6 radians/mm and slowly increases, thus generating cochlear models with successively more densely spaced reflection sites. The average filter bandwidths are plotted below in terms of 2π times inverse wavenumber; this quantity has units of length (mm) and is directly comparable to the wavelength of the TW at its peak.

Figure 4.9 summarizes the results of stability tests of four sets of two hundred different cochlear models, each with unique, spatially random gain distributions.

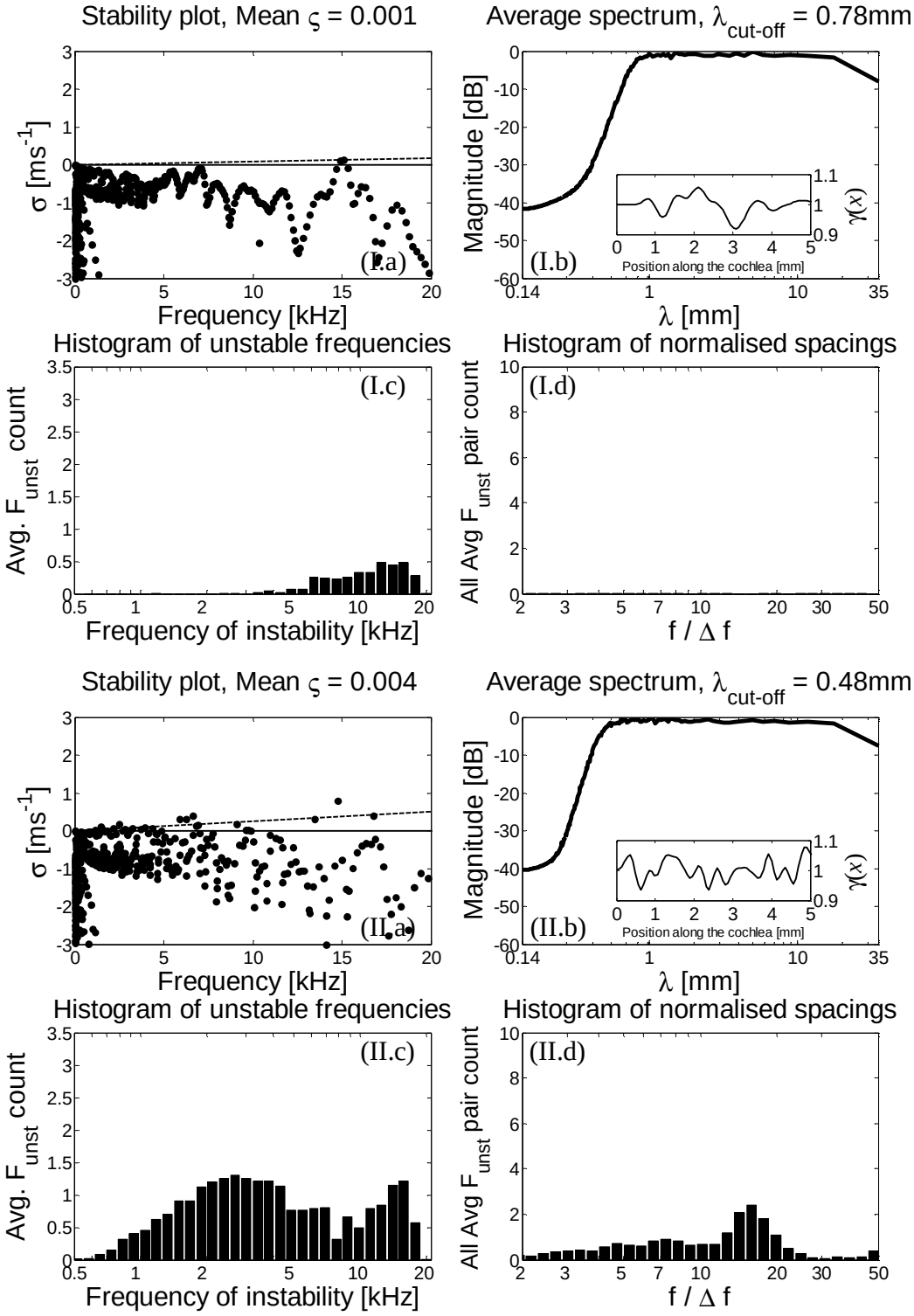
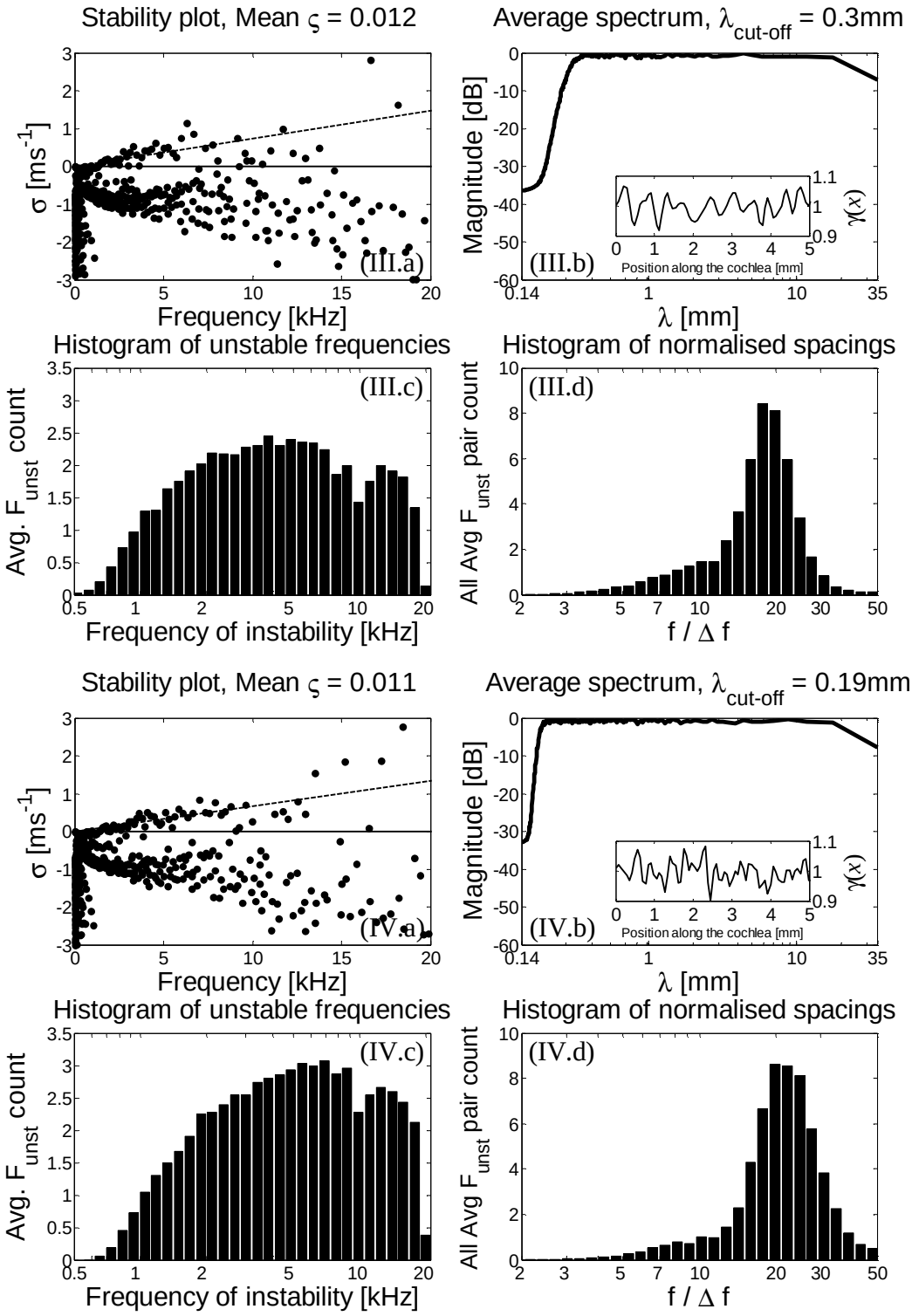


Figure 4.9.I-IV.a-d: The collected results from 4×200 cochlear models with randomly generated gain distributions. Each Roman numeral subset has been filtered with a different cut-off wavelength: (I.a-d) $\lambda_{cut\text{-}off}$ = 0.78 mm, (II.a-d) $\lambda_{cut\text{-}off}$ = 0.48 mm, (III.a-d) $\lambda_{cut\text{-}off}$ = 0.30 mm, (IV.a-d) $\lambda_{cut\text{-}off}$ = 0.19 mm. A peak-to-peak amplitude of 15% was applied to these gain distributions. *Caption continues on next page.*



(a) A characteristic stability plot taken from the set. The average undamping ratio for that single case, ς, is given and superimposed (dotted line). (b) Averaged inverse wavenumber spectrum of the gain; the first 5 mm of a characteristic gain distribution are inset. (c) Averaged histogram of all unstable frequencies per cochlea sorted in logarithmic frequency bins. (d) Averaged histogram of normalised spacings ($f/\Delta f$) per cochlea.

The four Roman numeral sets of Figure 4.9(.I – IV) have increasing high-wavenumber (decreasing low wavelength) cut-offs; this can be thought of as increasing the density of reflection sites in the cochlea. When the gain is smoothly varying with $\lambda_{cut\text{-}off}$ = 0.78 mm, as given in the I set, very few unstable frequencies are generated except near the base where the enhancement is the highest and most prone to instability. When $\lambda_{cut\text{-}off}$ is decreased to 0.48 mm, as given in set II, a significant number of instabilities appear at frequencies less than 5 kHz. The bin where the most instability is generated continues to increase in frequency as the cut-off wavelength is further decreased in Figure 4.9.III-IV. This is consistent with the results shown in the previous section.

When a given inverse wavenumber band becomes well-expressed in $\gamma(x)$, the corresponding region of the cochlea where λ_{peak} is twice the wavelength of the inhomogeneity becomes more prone to instability. Thus, the upward spread of the unstable frequencies with increasingly dense reflection sites is a direct result of the decreasing values of λ_{peak} toward the base. A distribution of $\gamma(x)$ is defined as 'dense' if its inverse wavenumber spectrum is well represented at half the λ_{peak} at the characteristic place of the frequency range of interest. Conversely, a distribution of $\gamma(x)$ is considered 'sparse' if this condition is not met.

Another related trend concerns the spacings of the unstable frequencies. In Figure 4.9.II.d, where $\lambda_{cut\text{-}off}$ = 0.48 mm, the peak in the spacings is $f/\Delta f \approx 15$ because most of the unstable frequencies fall within the region of 0.5 to 5 kHz. Here, the inverse wavenumber spectra of the inhomogeneities contain components at half λ_{peak} in the region $x >$ ~20 mm and are considered 'dense' for the frequencies less than several kHz; however, it is still 'sparse' for higher frequencies. As the cut-off wavelength is decreased, $\gamma(x)$ becomes 'dense' for more regions of the cochlea. These regions then express their predicted spacing as given in Figure 4.1.b; this shifts peaks in the spacings histograms in the (d) panels of Figure 4.9 toward successively higher values. It is possible to view this trend by plotting all of the spacings of adjacent unstable frequencies against the geometric average of the two frequencies, as shown in Figure 4.10.

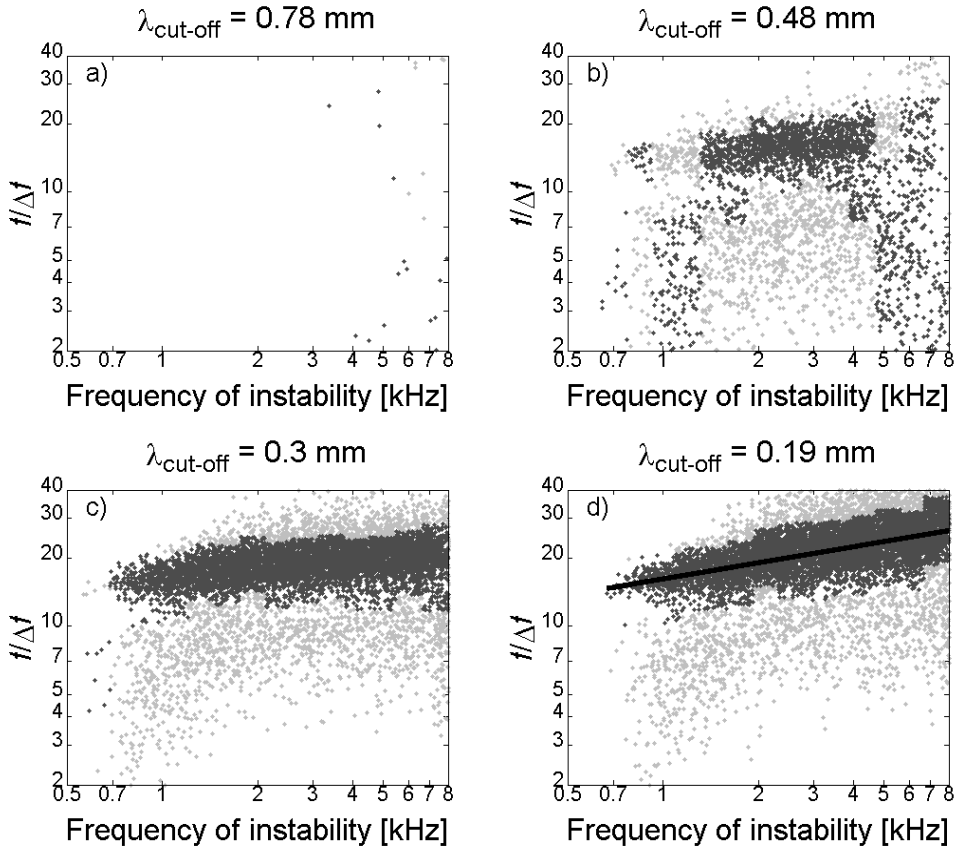


Figure 4.10.a-d: Plot of spacings between adjacent instabilities given various $\lambda_{cut\text{-}off}$ as a function of frequency for the results collected in Figure 4.9. The darkened dots represent spacings that fall within ± 1 standard deviation of the mode within 15 log-spaced bands. A trend line through the modes of each band is shown in panel (d).

Figure 4.10 shows that the spacings between instabilities are fairly widely spread apart when $\lambda_{cut\text{-}off}$ > approximately 0.48 mm, as suggested by Figure 4.1.a. There are wide spacings (low values of $f/\Delta f$) between instabilities at all frequencies in panel (b). When $\lambda_{cut\text{-}off}$ = 0.30 mm, as in Figure 4.10.c, a strong trend is visible throughout most of the frequency range. This trend is fully realized when the distribution of $\gamma(x)$ is 'dense' for all regions of the cochlea, as in panel (d), and appears to match the predictions of Figure 4.1.b when the CF for each position is calculated. The results of Figure 4.10.d qualitatively agree very well with the statistics of clinically measured SOAE spacings in humans, as presented by Shera (2003). However, there does appear to be a stronger bias in the linear model's results for very closely spaced instabilities, i.e. larger values of $f/\Delta f$, than in Figure 3 of (Shera, 2003); this is potentially reduced by the nonlinear interaction of adjacent SOAEs; this topic is discussed in the following subsection.

It should be noted that the reverse-transmission characteristics of the middle ear are similar to those of a band-pass filter centred at approximately 1 kHz, as shown for example in Figure A.15 in Appendix A. This would in theory reduce the number of SOAEs detected at frequencies far outside of the middle ear's pass-band. One would then expect that the commonly observed PMD between SOAE frequencies in humans is somewhat different from that observed here, as linear instabilities exist across a wide range of frequencies in the model. This is further discussed at the end of this chapter.

One final plot that also summarises a large number of linear stability simulations is Figure 4.11, which shows the average undamping ratio and number of instabilities in a cochlear model when varying both the amplitude of the perturbation in $\gamma(x)$ and its $\lambda_{cut\text{-}off}$.

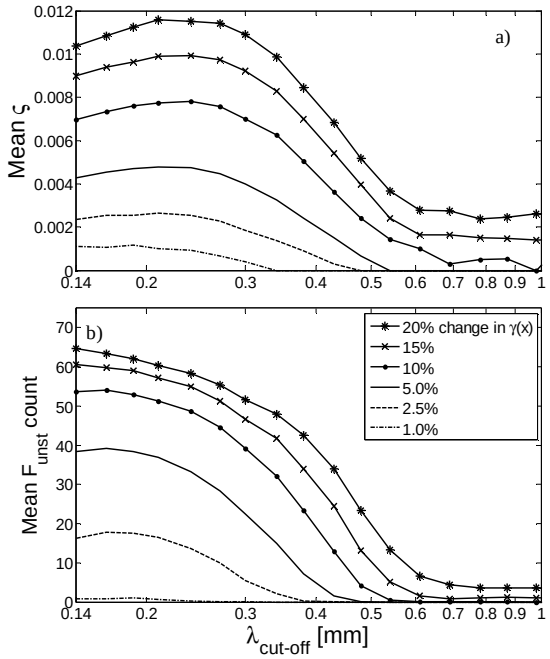


Figure 4.11.a-b: Average undamping ratio (a) and number of unstable frequencies per cochlear model (b) given variations in $\lambda_{cut\text{-}off}$ and the peak-to-peak variations in $\gamma(x)$.

The mean unstable frequency count and the mean undamping ratio, ς, vary directly with the amplitude of the variation in $\gamma(x)$. This result is consistent with the findings of Elliott *et al.* (2007). The statistics of the spacings of instabilities are largely independent of the exact form of the spatial variations, provided they have a significant component at the wavenumber corresponding to one half λ_{peak}. Peak-to-peak variations in $\gamma(x)$ as small as 1% can give rise to instabilities provided the distribution of feedback gain is 'dense.'

4.3 Nonlinear time domain simulations of unstable cochleae

Zweig and Shera's (1995) multiple-reflection theory of SOAE generation is capable of explaining many of the characteristics of spontaneous emissions in humans. It also provides a framework for understanding what is observed in the linear stability analyses of the cochlear model. However, the phenomenon this theory seeks to explain is nonlinear by nature; it is somewhat surprising that a linear theory of cochlear mechanics is able to accurately predict so many of its attributes.

This subsection seeks to begin to reconcile the linear theory of SOAE generation with nonlinear simulations of the cochlea in the time domain. In subsection 4.3.1, a step change in gain is implemented to study the evolution of a single unstable pole in time. Further complexity is added in subsection 4.3.2 where a small region of the cochlea is perturbed by random, 'dense' inhomogeneities in $\gamma(x)$. This dense set of reflection sites gives rise to multiple linearly unstable frequencies; the nonlinear simulation of this system in time shows how limit cycles can interact with one another.

The nonlinear simulations in this subsection are driven by a 100 μs-long click at 40 dB SPL in the blocked ear canal. The click is composed of a well-distributed range of frequency components that stimulates the entire cochlear model. This is necessary to briefly excite the unstable modes of these models, though the primary goal of these experiments is not to analyse the transients that are produced by this input. The details of the click stimulus are presented in Appendix A, and the model's response to clicks alone is investigated in the following chapter. In addition, the term 'steady state' is used in this chapter to describe responses that are examined at $t > 1000$ ms after the onset of the stimulus, by which time the transient has largely decayed away.

4.3.1 Step changes in $\gamma(x)$

A step change in $\gamma(x)$ is introduced which generates a single instability at 1.2135 kHz, as shown in Figure 4.12. When this system is stimulated with a click, a limit cycle is generated close to the unstable frequency. The steady state frequency of the nonlinear limit cycle is 1.2140 kHz; this value is within 0.04% of the linear prediction. The first 60 ms of the global response of this nonlinear model is illustrated in Figure 4.13 which shows how the initial wave following the click stimulus propagates from the base to the apex.

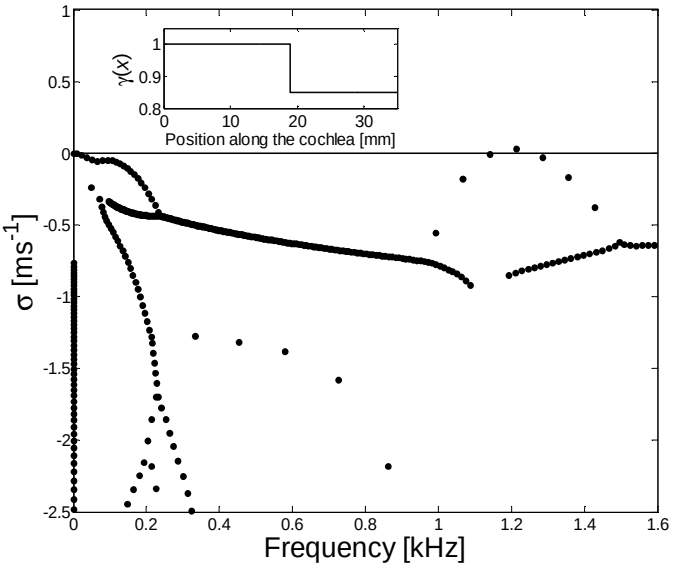


Figure 4.12.a-b: Stability plot of a model with $\gamma(x < 18.9 \text{ mm}) = 1$, $\gamma(x \geq 18.9 \text{ mm}) = 0.97$.

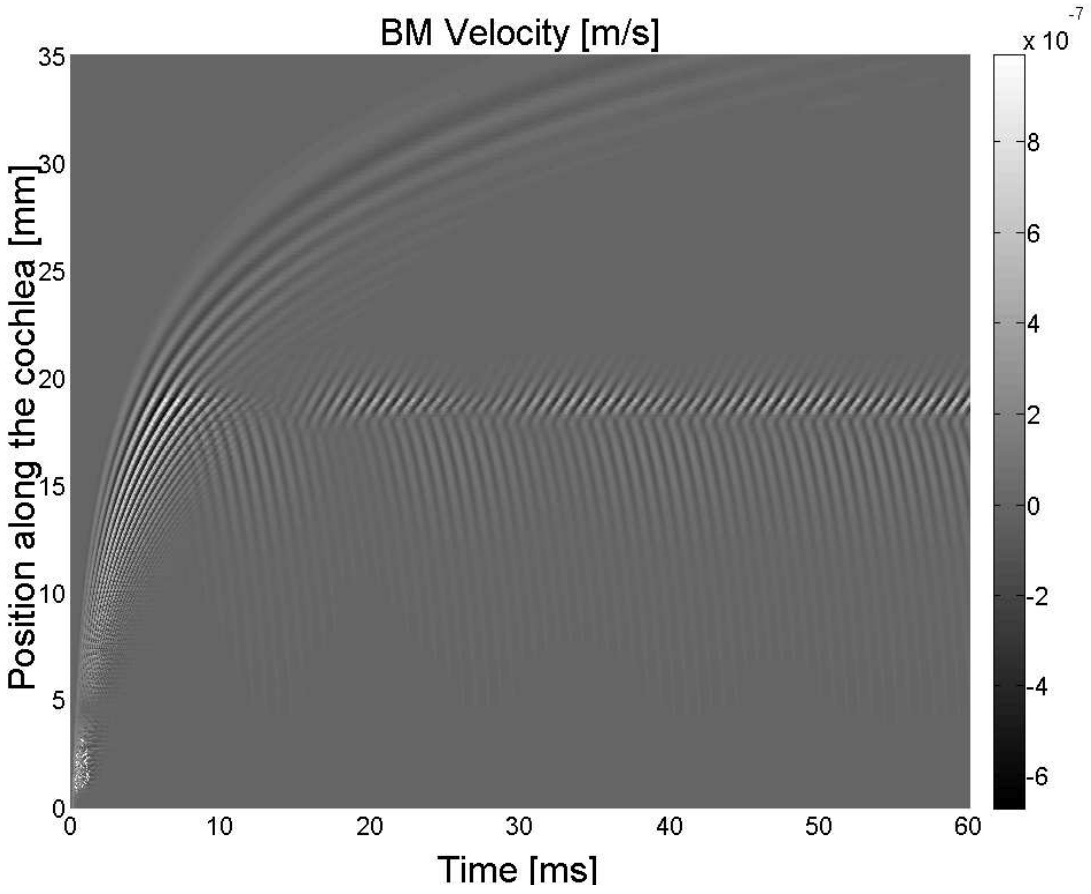


Figure 4.13.a-b: Mesh of results from a nonlinear unstable cochlea, simulated in the time domain. The stimulus applied was a 100 μs click.

When the transient wave reaches the step-discontinuity at approximately 7 ms, backward travelling waves are reflected from this site. The forward TWs in this mesh have a positive slope, whereas backward TWs have a negative slope.

Figure 4.14 plots the response of the BM at several locations; positions are plotted apically to basally in the three panels to facilitate comparison with the mesh plots. Figure 4.14.c shows the response of a point basal to the characteristic place of the instability. At this position, the transient click response reaches a peak at approximately 2 ms. At the characteristic place of the instability, as shown in Figure 4.14.b, the transient reaches a peak at approximately 6.5 ms. Finally, at a location apical of the best place, the initial transient reaches a peak at approximately 15 ms, as given in Figure 4.14.a. This increasing delay in the transient illustrates that the click is indeed propagating apically. Furthermore, the delays basal to the discontinuity in feedback gain are consistent with the predictions of Figure 2.21 where $\gamma = 1$.

One interesting feature of Figure 4.13 is that there are initially time frames of stronger and weaker activity near the location of the discontinuity. For instance, at x = 18.8 mm, there is a local maximum in the envelope of the response at approximately 21 ms, while there is a local minimum in the envelope of the response at approximately 28 ms. The peaks and dips in the envelope of the response at the characteristic place of the instability are due to the sloshing about of reflected energy from the original stimulus. It takes approximately 7 ms for the energy of the click to propagate from the base to the discontinuity in $\gamma(x)$ located at x = 18.8 mm. A portion of its energy is subsequently reflected back toward the base during the next 7 ms. When this backward TW encounters the middle ear, it is again partly reflected back toward the discontinuity. After another 7 ms, this initial wavelet, now twice-reflected, again reaches the discontinuity. This explains why the peaks in the envelope of the response fall at odd multiples of the cochlear delay to the location of the discontinuity: $\tau(x = 18.8 \text{ mm}) \approx 7$ ms, $3\tau(x = 18.8 \text{ mm}) \approx 21$ ms, etc.

While the transient response dies away, oscillations at the unstable frequency add coherently over successive reflections between the base and the step-discontinuity, as shown in Figure 4.14. Oscillations gradually settle into a fixed-amplitude limit cycle that peaks close to its characteristic place, the response of which is plotted in Figure 4.14.b. Low-level oscillations of these reflected TWs (at the unstable frequency) are visible at the location basal to the discontinuity in Figure 4.14.c for $t > 10$ ms. However, at $x > 18.9$

mm, the response is quickly extinguished and no motion is visible after the initial transient in Figure 4.14.a.

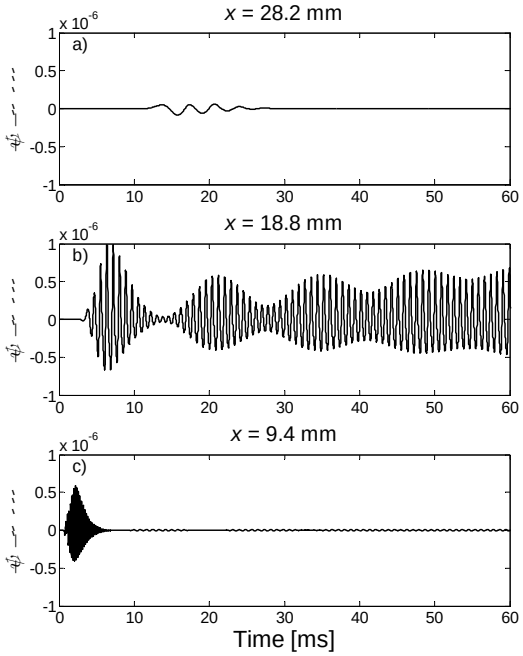


Figure 4.14.a-c: The first 60 ms of BM velocity at three locations in the unstable cochlear model shown in Figure 4.12: a) x = 28.2 mm; b) x = 18.8 mm; c) x = 9.4 mm.

The time-evolution of the frequency components in the pressure at the base of the cochlea is plotted in Figure 4.15 over several different time frames. There is a clear influence here from the frequencies of the near-unstable poles in the stability plot, Figure 4.12.

Blunt peaks are visible in the pressure spectrum of Figure 4.15.a at frequencies corresponding to the near-unstable poles as well as the single unstable frequency. The levels of these peaks also seem well-correlated with the relative magnitudes of the real parts of the poles, σ_i. This is consistent with the calculated linear transient response of a system with damped modes. Such a linear system would include spectral components at each of the natural frequencies of the system. The decay of these components is determined by the corresponding values of σ_i. For this nonlinear system, harmonic distortion is visible at the second and third harmonics of the fundamentals. As later time frames are examined, only the response at the linearly unstable frequency and its harmonics persist. Notice that the spectral resolution of the panels improves as longer time windows are analysed at later time frames. The time-variation of several frequency components in the pressure response is compared in Figure 4.16.

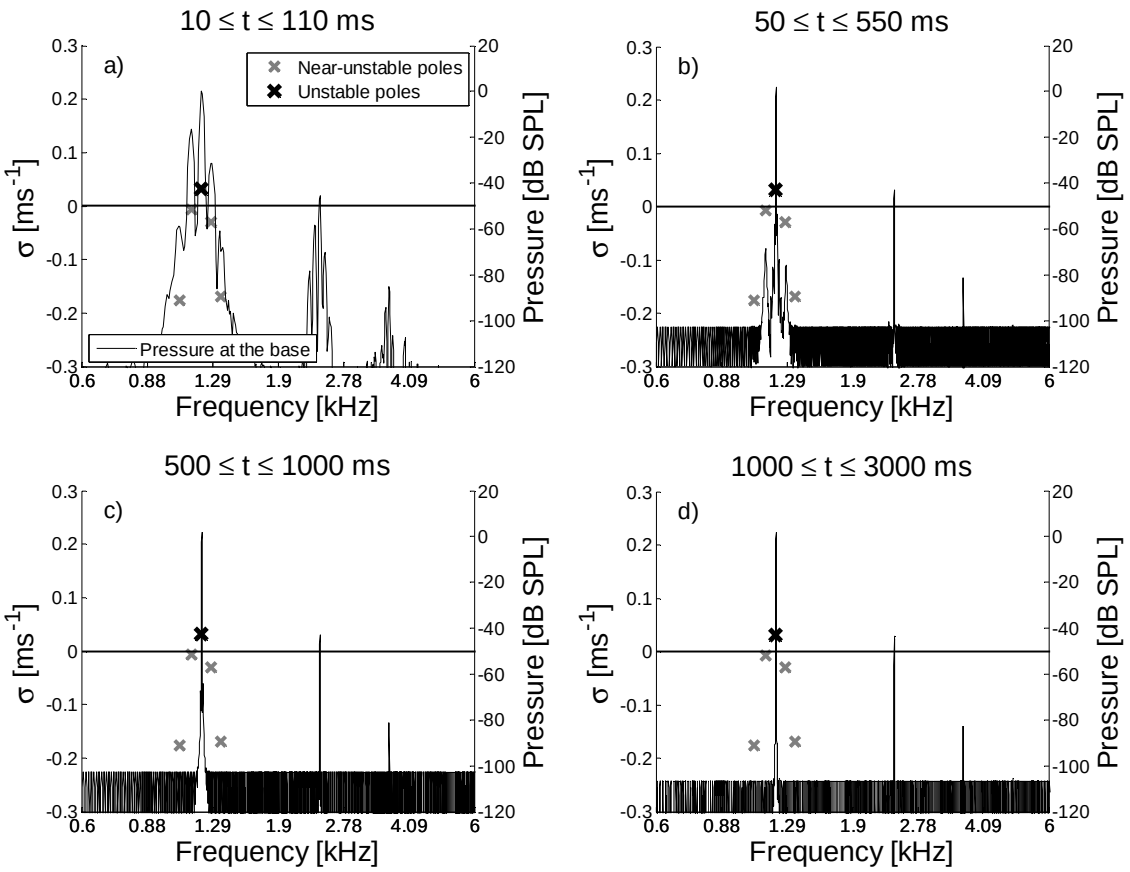


Figure 4.15.a-d: Superimposed stability plots of linear system and the spectrum of the pressure at the base of the nonlinear cochlea given four time windows: a) $10 \leq t \leq 110$ ms; b) $50 \leq t \leq 550$ ms; c) $500 \leq t \leq 1500$ ms; d) $1000 \leq t \leq 3000$ ms. Unstable poles are plot with dark '**x**'s; stable poles are plot with light '**x**'s.

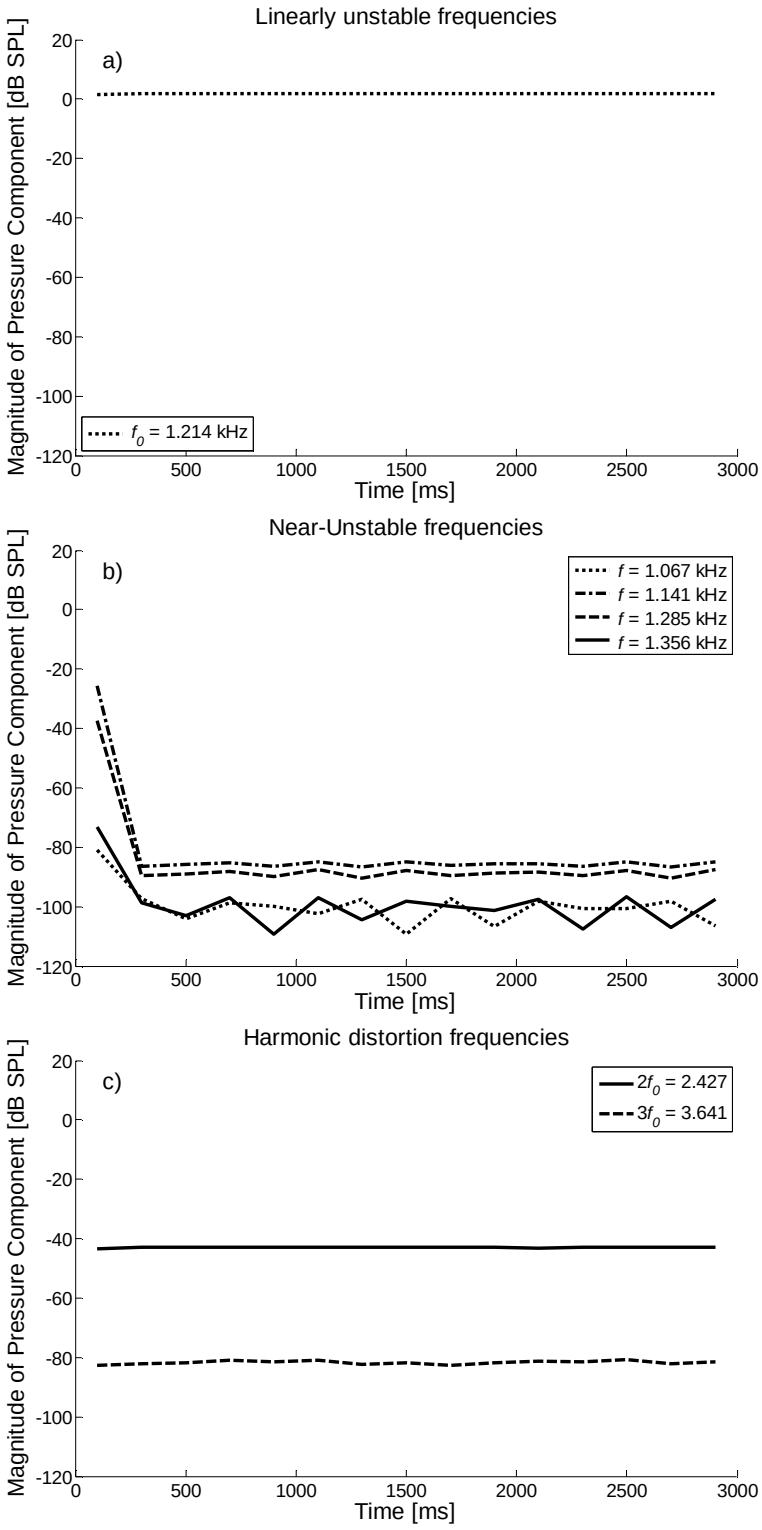


Figure 4.16.a-c: Variation of the magnitudes of various frequency components of the pressure at the base with time in an unstable cochlea. Linearly unstable and near-unstable frequencies are shown in panels (a) and (b), respectively, while harmonic distortion frequencies are shown in panel (c). Every curve consists of 15 data points, where each value represents the DFT of 200 ms of data with no overlap between adjacent windows.

It is clear from Figure 4.16 that the magnitude of the unstable frequency and its harmonics remain relatively constant with time. However, the near-unstable frequencies decay away as predicted by linear stability analysis. Though the relative initial levels of the near-unstable frequencies correlate well with the σ_i of their respective poles, the final values of the f = 1.285 and 1.114 kHz components are somewhat stronger than the other three plotted frequencies. Nevertheless, all of the near-unstable components are at least 90 dB below that of the primary unstable frequency at the end of the simulation.

The steady state magnitudes and phases of the primary and its first two harmonics are plot as a function of position in Figure 4.17.

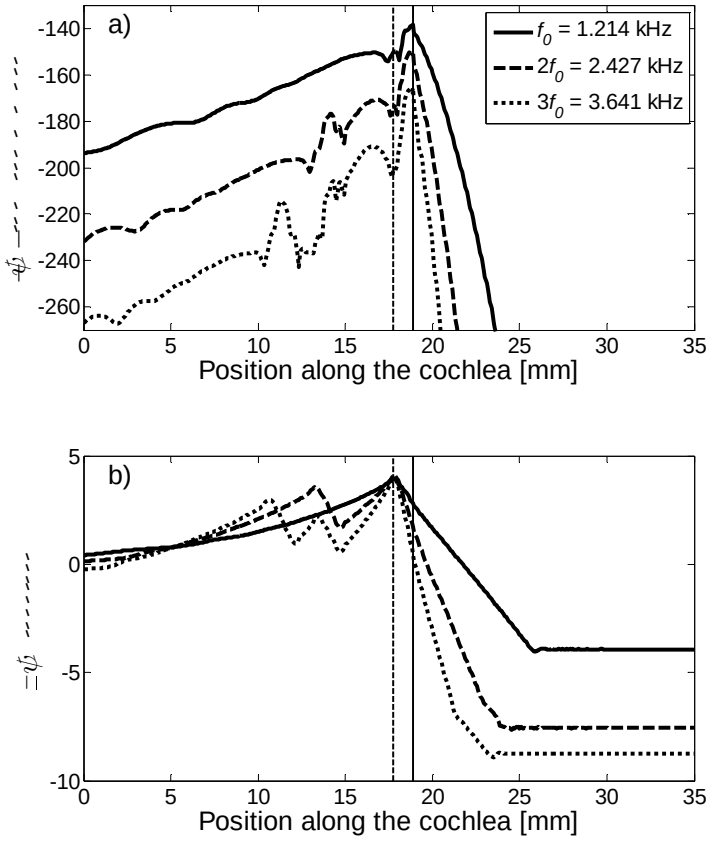


Figure 4.17.a-b: The magnitude (a) and phase (b) of the BM velocity at the unstable frequency and its first two harmonics calculated as a function of position. The fundamental response is plotted as a solid line, whereas the $2f_0$ and $3f_0$ components are plotted as a dashed and dotted lines, respectively. A solid vertical line marks the location of the maximum magnitude response of the fundamental, whereas a dashed vertical line marks the location of the maximum phase response of the fundamental.

The magnitudes of the various components in Figure 4.17.a indicate that the $2f_0$ and $3f_0$ harmonics of the unstable frequency both have local maxima at their characteristic places (at approximately 14 mm and 11.5 mm, respectively). However, the absolute maximum peak of each frequency is close to the characteristic place of the fundamental. The magnitudes of the BM harmonic distortion components at the base of the cochlea are approximately 40 and 80 dB below the fundamental in Figure 4.17; this is confirmed by the pressure response at the base in Figure 4.16.c. At x = 18.9 mm, however, the peaks of the three frequencies fall within a range of only 27 dB. This shows that the fundamental has been amplified less than the harmonics, perhaps due to self-suppression. The phases of Figure 4.17.b can reveal further details of the response, as discussed at the end of this chapter.

4.3.2 Random variation in γ(x)

As illustrated in the simulations of a single unstable pole, a saturation nonlinearity in the feedback loop is necessary to limit the amplitude of the oscillations of the linear instability. The unstable frequency continued to ring indefinitely after the near-unstable frequencies died away. In this subsection, further complexity is considered by examining the nonlinear interaction of several linear instabilities in the cochlear model.

Only a restricted region of the cochlea is perturbed with a 'dense' distribution of inhomogeneities, so that the interaction of only a relatively small number of unstable poles can be observed. This is accomplished by windowing a distribution of $\gamma(x)$ within the range of $16.5 < x < 23.5$ mm, as shown in Figure 4.18. A 3.5 mm-long Hanning window was generated and extended by filling its centre with 3.5 mm of ones; zeros were padded outside this range to extinguish the remaining variations in gain outside the window. The stability of a cochlear model that incorporates this windowed-perturbed gain distribution is presented in Figure 4.19. The perturbed poles fall within a range of ~0.7 to 1.5 kHz, which corresponds to the characteristic frequencies within the spatial window function.

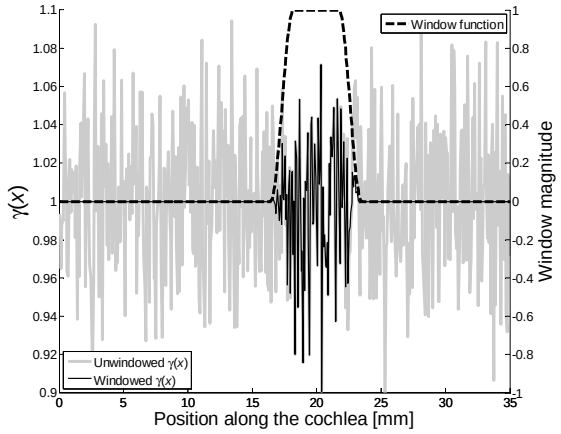


Figure 4.18: A distribution of $\gamma(x)$ is shown in thick grey. A 7-mm extended Hanning window is applied, centred at x = 19 mm. The resultant windowed distribution of $\gamma(x)$ is shown in thin black.

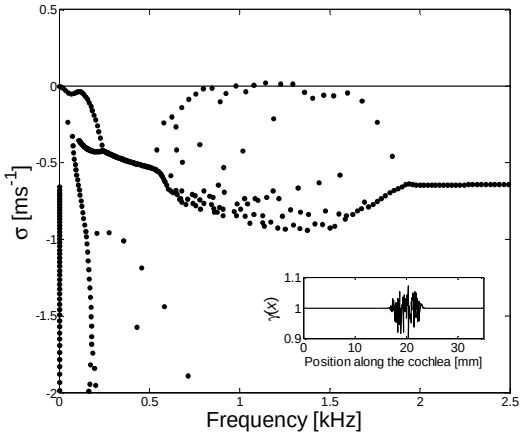


Figure 4.19: Stability of a cochlear model given the windowed-perturbed gain distribution as presented in Figure 4.18.

Five distinct instabilities are generated by this windowed $\gamma(x)$, located at f = [0.979, 1.080, 1.145, 1.229, 1.296] kHz. The undamping ratio of these poles is ς = [0.004, 0.009, 0.030, 0.020, 0.020], respectively. There are also a number of near-unstable poles, located at f = [0.800, 0.848, 0.923, 1.033, 1.359, 1.597] kHz.

A mesh of the first 60 ms of BM velocity is shown in Figure 4.20; a 100 μs click was the only stimulus. Figure 4.20 shows that the greatest BM activity is restricted to the perturbed region after the initial transient decays away; this is centred at $x \approx 19$ mm. Again, backward travelling waves are visible from approximately 8 ms onward from the basal edge of the perturbed area.

This response is further clarified in Figure 4.21, which shows the BM velocity at locations basal to-, within-, and apical of the perturbed region. As with the simulation of a single unstable pole in the previous subsection, only locations near the inhomogeneities exhibit much activity after the initial stimulus has decayed away. However, the activity in this randomly perturbed region, as shown in Figure 4.21.b, is clearly more complicated than the equivalent plot of the step-perturbed response, as previously given in Figure 4.14.b.

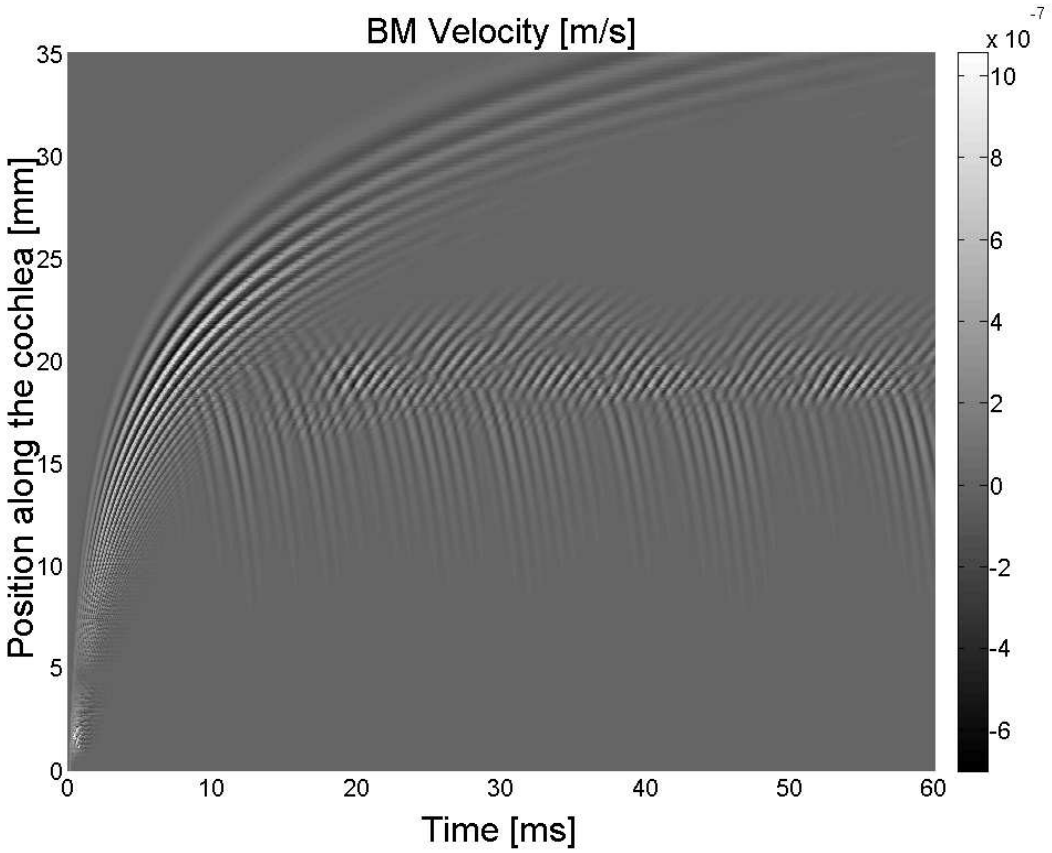


Figure 4.20: Mesh of BM velocity in the first 60 ms of a nonlinear simulation of an unstable cochlear model; the stability of this system is shown in Figure 4.19.

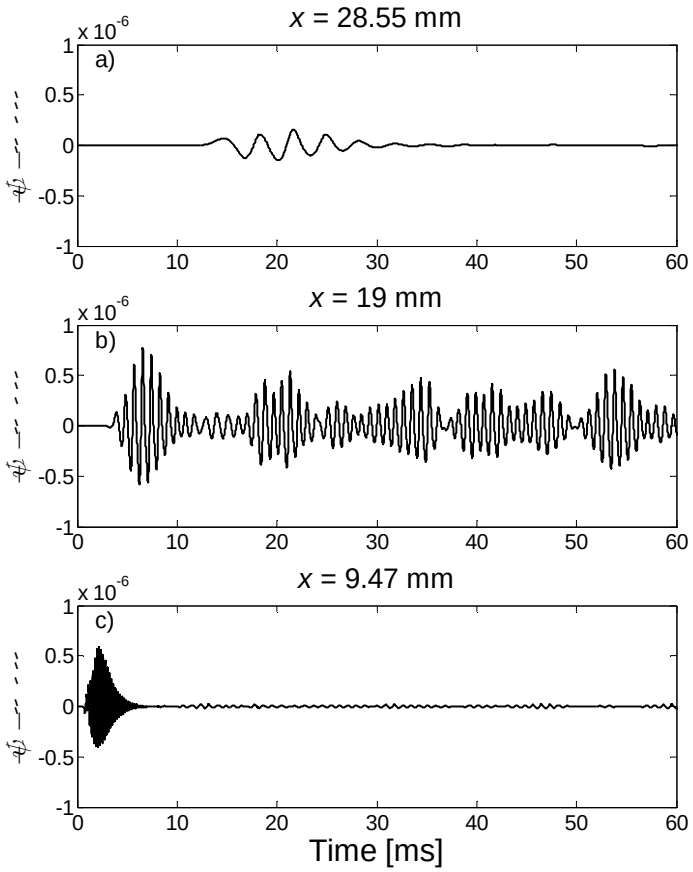


Figure 4.21: The first 60 ms of BM velocity at three different locations following a click stimulus at the base: a) x = 9.47 mm; b) x = 19 mm; c) x = 28.55 mm.

The spectrum of the pressure at the base of the cochlea is calculated from the last 2000 ms of data in the 3000 ms-long simulation to avoid any contamination from the initial stimulus. This is plotted simultaneously with the stability plot of the system in Figure 4.22. Unless otherwise noted, a Blackman window has been applied to all of the spectral results in this section; this form was chosen for its sharp attenuation of sidebands, though other window functions with similarly sharp cut-off characteristics produced similar results.

Figure 4.22 shows that there are more than 40 distinguishable peaks in the spectrum of the pressure at the base, with amplitudes ranging from near 0 dB SPL down to approximately -105 dB SPL in this simulation.

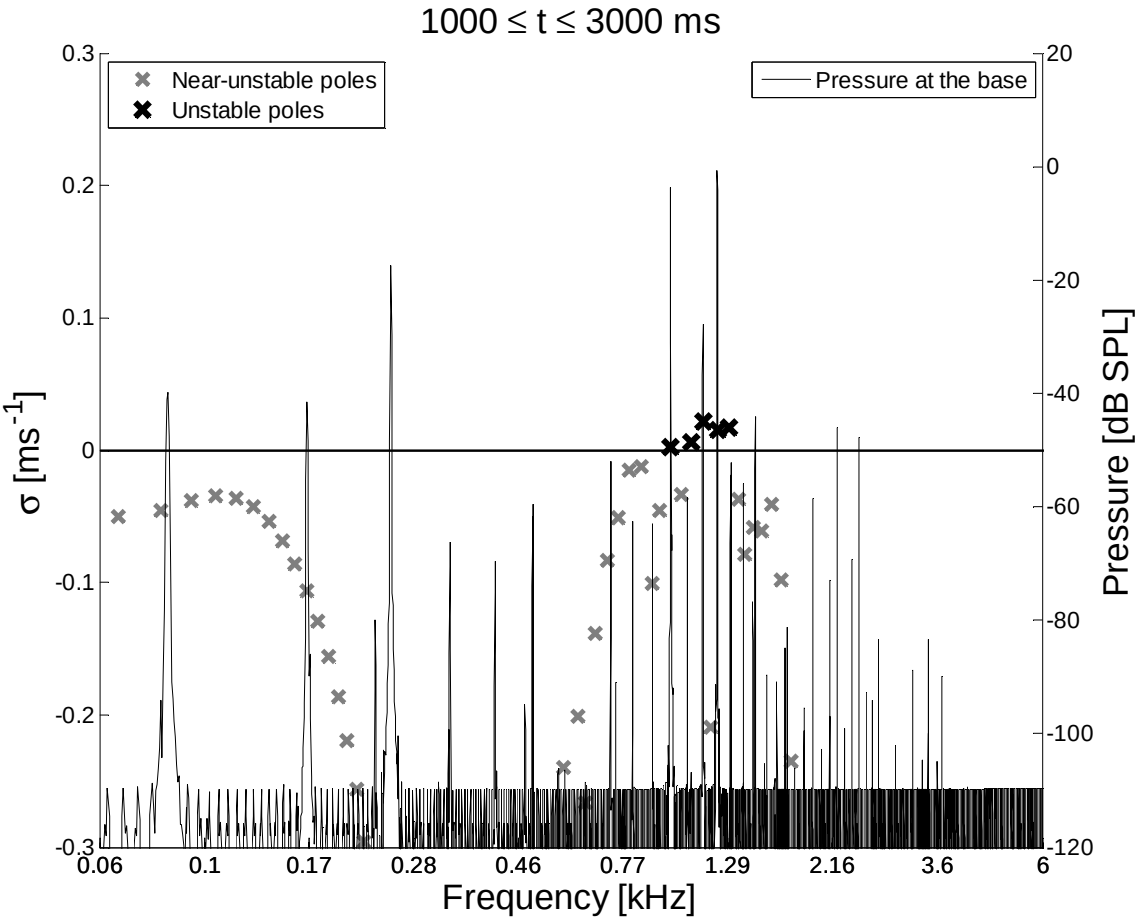


Figure 4.22: Simultaneous plot of linear system stability and the pressure spectrum at the base of the cochlea for $1000 \leq t \leq 3000$ ms. Unstable poles are plot with dark '**x**'s; stable poles are plot with light '**x**'s.

The range of these frequency components extends from 0.083 kHz all the way up to 3.7 kHz. This wide range of activity is the result of the nonlinear interactions of only five linearly unstable poles. The 'new' frequencies that were not predicted by the linear stability analysis can be attributed to both harmonic and intermodulation distortion. For instance, the peak at 2.455 kHz and -48 dB SPL appears to be the second harmonic of the linearly unstable limit cycle at 1.227 kHz. Similarly, the lowest frequency peak at 0.084 kHz and -40 dB SPL appears to be a difference tone resulting from the limit cycles of the two linearly unstable frequencies at 1.227 kHz and 1.143 kHz. What is also interesting about Figure 4.22 is that only three of the five linearly unstable frequencies are expressed in the last 2000 ms of the simulation. The limit cycle frequencies that do persist are all within 0.1% of the linearly predicted unstable frequencies.

In order to understand the time-evolution of the various frequency components, Figure 4.23 shows the spectrum of the pressure at the base, calculated over four different time frames. Only a tight range of frequencies near the linearly unstable frequencies are shown for clarity. In the earliest frame, shown in panel (a), there are peaks at each fully-unstable and near-unstable frequency as predicted by linear stability analysis. As one might intuitively expect from the response of linear systems, the magnitude of each peak is directly related to the magnitude of the real part of the corresponding pole. As the time window is shifted later in time and further away from the initial stimulus, some of these initial peaks begin to recede in amplitude. Almost all of the (linearly) near-unstable frequencies have fallen to the noise floor by panel (d). In addition, other distortion-related frequencies begin to rise at later time frames as well.

The variation of the amplitudes of different components of the pressure response is charted as a function of time in Figure 4.24. Three separate panels show the DFT magnitudes of the linearly unstable, linearly near-unstable, and a number of predicted distortion product frequencies.

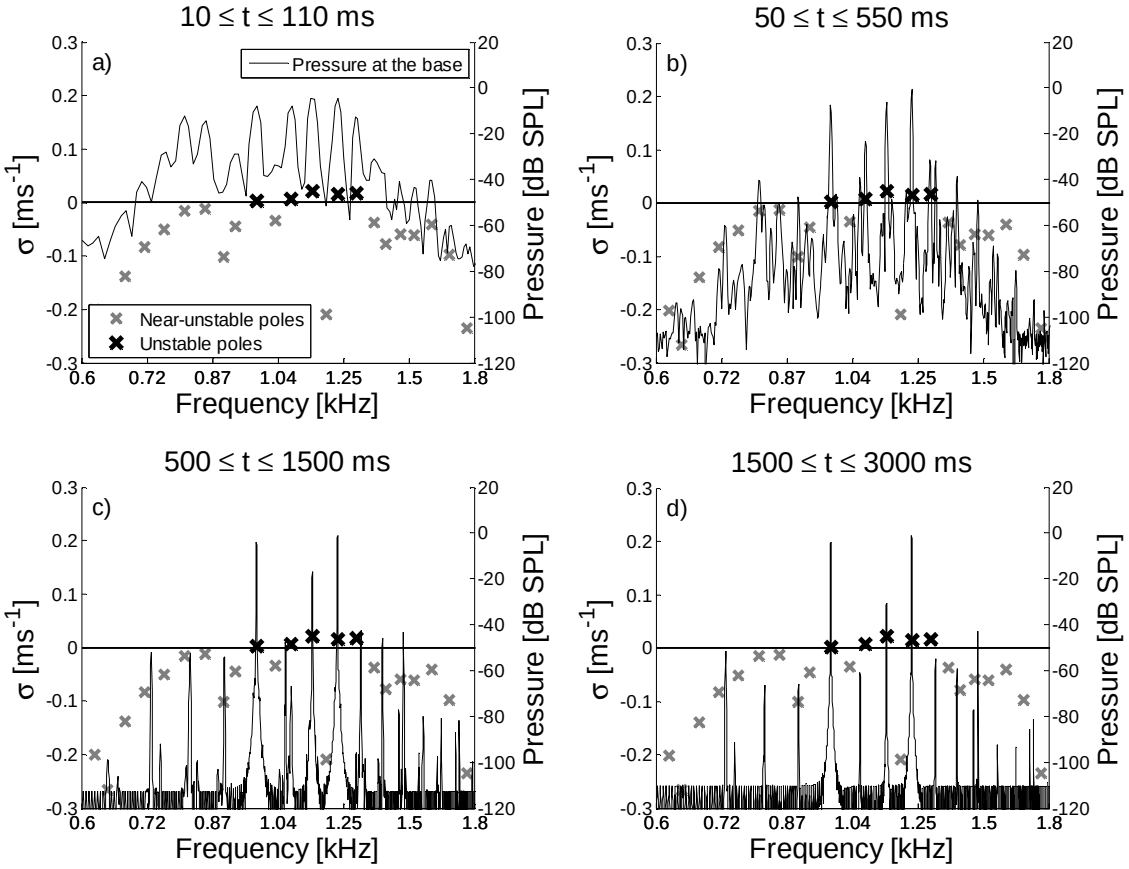


Figure 4.23.a-d: Simultaneous plot of linear system stability and the spectrum of the pressure at the base of the cochlea given four time windows: a) $10 \leq t \leq 110$ ms; b) $50 \leq t \leq 550$ ms; c) $500 \leq t \leq 1500$ ms; d) $1500 \leq t \leq 3000$ ms. Unstable poles are plot with dark '**x**'s; stable poles are plot with light '**x**'s.

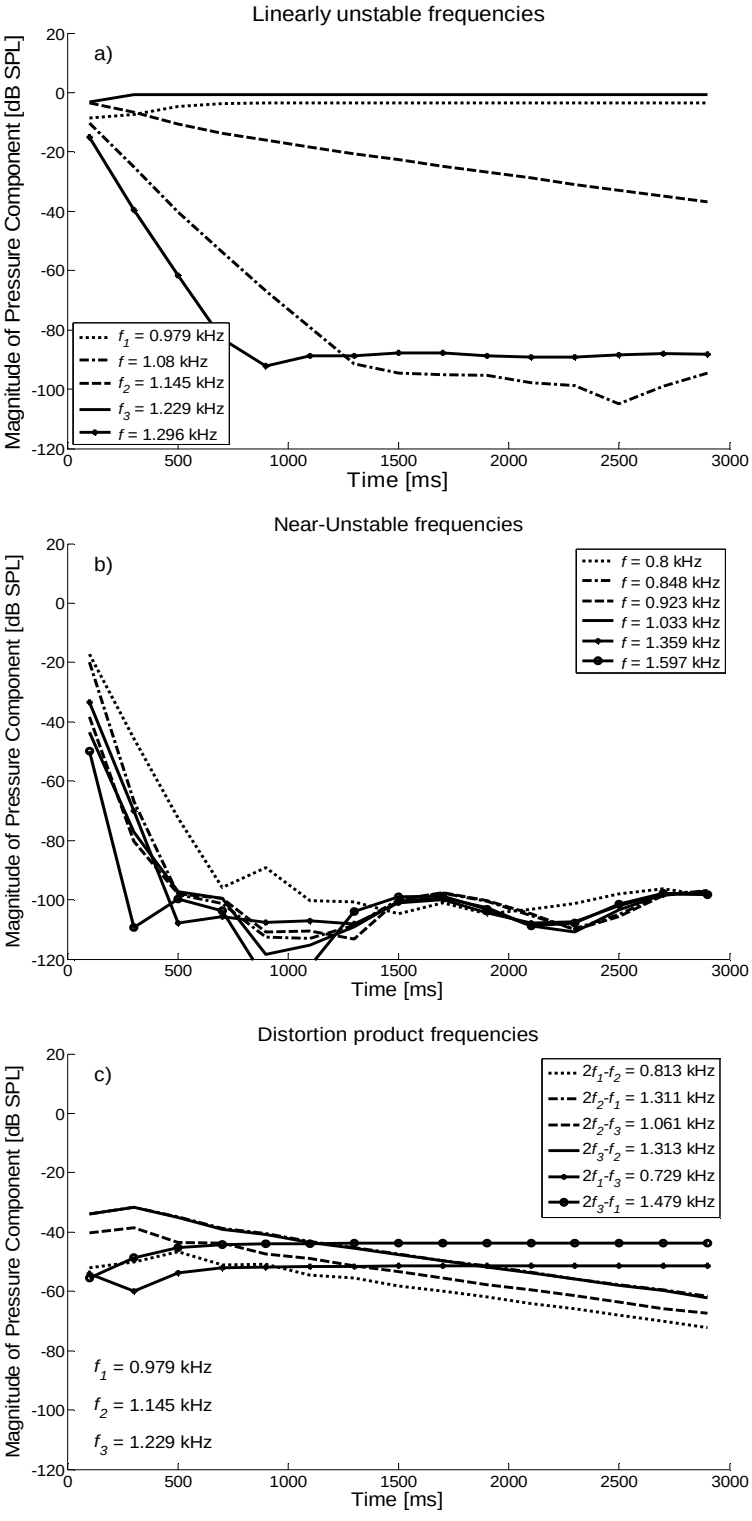


Figure 4.24.a-c: Variation of the magnitudes of various frequency components of the pressure at the base with time in an unstable cochlea. Linearly unstable and near-unstable frequencies are shown in panels (a) and (b), respectively, while distortion product frequencies are shown in panel (c). Every curve consists of 15 data points, where each value represents the DFT of 200 ms of data with no overlap between adjacent windows.

Figure 4.24.b illustrates the expected transient behaviour of the stable modes, as the magnitudes of the near-unstable frequencies fall to a level to near the noise floor approximately 500 ms after the initial stimulus. The initial rate of decay appears to be approximately -140 dB/s for each of the frequencies shown.

The magnitude of the frequency components that were determined to be unstable by linear stability analysis are plotted in Figure 4.24.a as a function of time. The responses of f = 1.229 kHz and f = 0.979 kHz increase slightly in amplitude during the first 500 ms, but then reach reasonably constant levels of -0.8 and -3.5 dB SPL, respectively. However, the remaining three linearly unstable frequencies decay away at various rates. For instance, f = 1.296 kHz and f = 1.080 kHz initially decay at approximately -100 dB/s and -70 dB/s, while f = 1.145 kHz recedes much more slowly at approximately -10 dB/s. It is clear that nonlinear suppression is taking place between the unstable tones, contrary to the linear analysis; the observations made here are discussed further at the end of the chapter.

The magnitudes of a number of commonly observed distortion products which result from three assumed primaries are given in Figure 4.24.c. In addition to the most commonly studied DPOAE, the cubic distortion product ($2f_l - f_h$), one other nearby DPOAE ($2f_h - f_l$) is examined for to its proximity to the primaries. A general notation of f_l and f_h, corresponding to the frequencies of the lower tone and the higher tone, is adopted above to avoid confusion with the notation for the selected primaries. The primaries chosen are the three linearly unstable frequencies that persist in amplitude: f_1 = 0.979 kHz, f_2 = 1.145 kHz and f_3 = 1.229 kHz. The magnitudes of the distortion products at $2f_1 - f_3$ and $2f_3 - f_1$ mirror the growth of the two primaries at f_1 and f_3, just as the magnitudes of the other four distortion products show slow decay, in a manner similar to f_2. Note however that decay rates of these distortion products are somewhat less steep than that of f_2; this is perhaps because the amplitude of the other primaries are stable.

One of the salient features of SOAEs, however, is the distribution of spacings between unstable frequencies. To examine the log-normalised spacings between adjacent limit cycles of this nonlinear simulation, an arbitrary threshold was set at -65 dB below the strongest instability to choose frequencies for analysis. The selected limit cycles and the resultant $f/\Delta f$ spacings are shown in Figure 4.25.

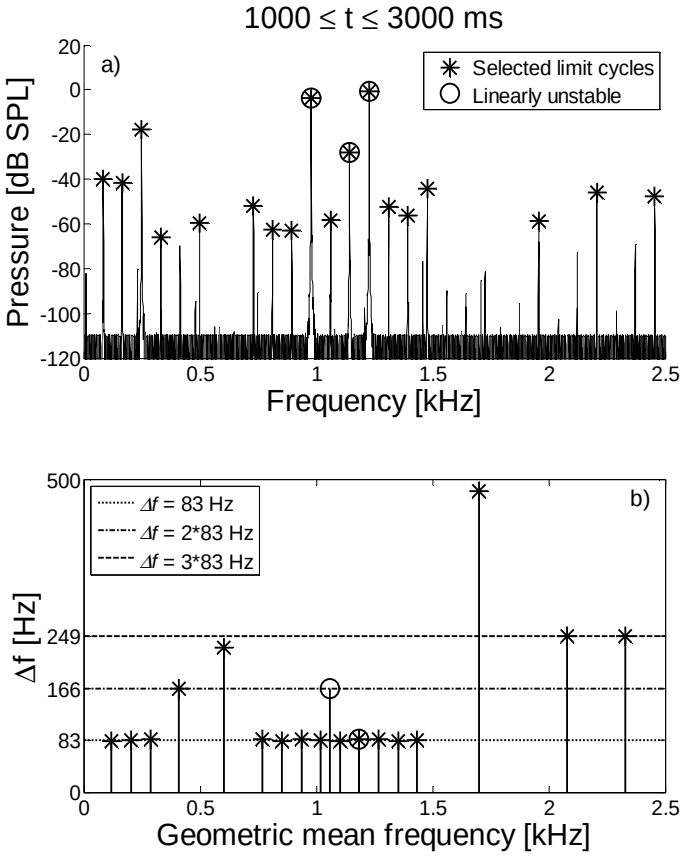


Figure 4.25.a-b: The pressure spectrum in panel (a) displays a number of selected limit cycles; these frequencies are used to compute distances between the instabilities, as shown in panel (b). Selected frequencies are indicated by a star (✳), and linearly unstable frequencies and spacings are denoted by a circle (O). The frequency axis of panel (b) represents the geometric mean of the two adjacent limit cycle frequencies. Dotted, dot-dashed and dashed horizontal lines are drawn at Δf = [83, 2*83, 3*83] Hz, respectively.

The results of Figure 4.25 are plotted on a linear frequency axis in order to emphasize the spectrally periodic nature of the limit cycles. These results are considered in subsection 4.4.2.

4.4 Discussion

The investigations presented in this chapter represent an attempt to understand the mechanisms that give rise to SOAEs. The approach taken here assumes that an unstable nonlinear system that saturates is capable of generating limit cycle oscillations similar to SOAEs. Thus, the linear stability of the cochlear model was a natural starting point for comparing theoretical predictions to simulations.

4.4.1 Linear results

The findings of section 4.2, based on a linear numerical model of the human cochlea, are consistent with the multiple-reflection theory of Zweig and Shera (1995). The state space formulation is able to predict the frequencies at which a linear, active cochlear model will become unstable. Direct measurements in animals have shown that spontaneous basilar membrane oscillation is associated with SOAEs in the ear canal (Nuttall *et al.*, 2004). Consequently, comparisons are drawn in this paper between measured SOAE characteristics and the instabilities generated in the cochlear model. However, it is worth highlighting the similarities and differences between measured data and these simulation results.

This model predicts that instabilities exist all along the cochlea and across a wide range of frequencies, given a dense array of inhomogeneities in the cochlea. In contrast, SOAEs in normal-hearing individuals are only routinely detected between 0.5 kHz to 6 kHz (Probst *et al.*, 1991). Even if instabilities exist in all regions along the average human cochlea, however, it is likely that only a subset of these will be detected in the ear canal. It is believed that the inefficient reverse-transmission characteristics of the middle ear hinder the detection of SOAEs outside of its best transmissibility range, given its steep drop-off above and below its ~1 kHz resonance of approximately -40dB per decade. The limited bandwidth of normally-detected SOAEs is also potentially reduced by physiological noise and the current limitations of sensor technology. Just as improved measurement techniques have revealed increasingly sharp active BM enhancement through the years, refinements in recording technique have exposed a higher prevalence of SOAEs in more recent studies (Probst *et al.*, 1991; Penner and Zhang, 1997).

The average number of unstable frequencies shown in Figure 4.11 for a ‘dense’ 5% peak-to-peak variation in gain is similar to the maximum number of emissions detected in a single ear, some in excess of 30 SOAEs (Talmadge *et al.*, 1993). It has been shown that the level and number of instabilities in the state space model depend on the amplitude of the variations in BM impedance and the spatial density of the inhomogeneities. For instance, the current linear model predicts a distribution of unstable frequency spacings that is similar to physiologically compiled data in several respects.

The spacings between linear instabilities in this model can match the observed variation in SOAE spacings with frequency as given by Shera (2003); however, this is only

realised when the wavenumber distribution of $\gamma(x)$ is 'dense.' For instance, the spacing results presented in Figure 4.9.IV.d and Figure 4.10.d are consistent with the model predictions shown in Figure 4.1.b. The peak of the normalised spacings in Figure 4.9.IV.d is located at a value of $f/\Delta f$ = 20, which is somewhat more closely spaced than the commonly-observed PMD. However, the peak in $f/\Delta f \approx 15$ in Figure 4.9.I.d which represents the situation where the distribution of $\gamma(x)$ is 'dense' only for the lower frequency regions of the cochlea (f_{cf} < 5 kHz). This is consistent with the results in the previous subsection given sinusoidal distributions of $\gamma(x)$ with longer wavelengths.

When the current understanding regarding hearing sensitivity, the various forms of OAEs and pathology are combined, a convincing picture regarding the generation of SOAEs begins to evolve. As many authors have pointed out, SOAEs in humans appear to be a natural by-product of the species' sharply tuned sense of hearing. Normal hearing individuals that do not exhibit SOAEs typically have an audiogram which underperforms those with SOAEs by approximately 3 dB in the standard 1-6 kHz range (McFadden and Mishra, 1993). Pélanová *et al.* (2007) also reported that the high-frequency audiogram of normal-hearing children without SOAEs underperformed those with SOAEs by approximately 5 dB through the 10-16 kHz range. In the 'laser-cochlea' theory of OAE generation, it is the portion of the cochlea basal to the reflection site that is crucial to sustaining the limit cycle oscillation. If the losses in this region are not overcome by the active enhancement provided by the outer hair cells, no spontaneous emission can occur. This is exemplified by the two stability plots given in Figure 3.8; a step down in gain from γ = 1 causes instability, but a step up in gain to γ = 1 at the same location maintains stability.

4.4.2 Nonlinear results

The nonlinear results of this chapter demonstrate that a great deal of complexity can arise from just a few linear instabilities. The nonlinear time domain simulations showed that linear reflection initially dominates the response following a click stimulus. For instance, in both Figure 4.15.a and Figure 4.23.a, there are blunt peaks in the basal pressure spectrum at all frequencies where poles have moved toward or beyond the boundary of stability. As the response settles, the near-unstable frequencies decay away; however, in the simulation of a model with a single instability, the unstable frequency persists and

stabilises into a limit cycle oscillation as shown in Figure 4.16. As a result of the limit cycle, harmonic distortion is also generated at $2f_0$ and $3f_0$, where f_0 is the frequency of the instability. The dynamics of the fundamental response are the simplest to describe so they are discussed first.

It is instructive to consider a one-dimensional backward-travelling wave in a homogeneous medium of unit amplitude that can be written as

$$y(t) = \cos(\omega t + \kappa_{\mathrm{TW}} x), \tag{4.6}$$

where ω is the angular frequency, t is time, x is position, and κ_{TW} is the wavenumber. The wavenumber can be expressed as

$$\kappa_{\mathrm{TW}} = \frac{2\pi}{\lambda_{\mathrm{TW}}} = \frac{\partial \phi}{\partial x}. \tag{4.7}$$

Thus, if the slope of the phase with respect to position is positive, then the TW is indeed backward-travelling. However, if $\partial\phi/\partial x$ is negative, then (4.6) represents a forward-travelling wave. Similarly, if the $\partial\phi/\partial x$ is locally zero, then (4.6) ceases to become a function of position and the wave no longer propagates in this region. In addition, the wave velocity, or phase speed, is related to the wavenumber by

$$c_{\mathrm{TW}} = \frac{\omega}{\kappa_{\mathrm{TW}}}. \tag{4.8}$$

The wave velocity of the TW can be calculated from the variation of its phase with position, as described by equations (4.7) and (4.8).

Returning to the results presented in Section 4.3, the direction of the dominant TW can be determined, assuming the wavenumber is changing slowly in a given region. Figure 4.26 plots the wave velocity of the three frequency components of the BM response in Figure 4.17. A positive velocity in Figure 4.26 indicates forward travel, whereas a negative velocity indicates backward travel. There are regions where the dominant TW is forward-propagating, and regions where it is backward-propagating. There are also positions where the TW is 'standing,' and beyond ~25 mm the TWs are extinguished. The response of the TW at the f_0 appears to change directions once, whereas the $2f_0$ and $3f_0$ TWs appear to change directions three and five times, respectively.

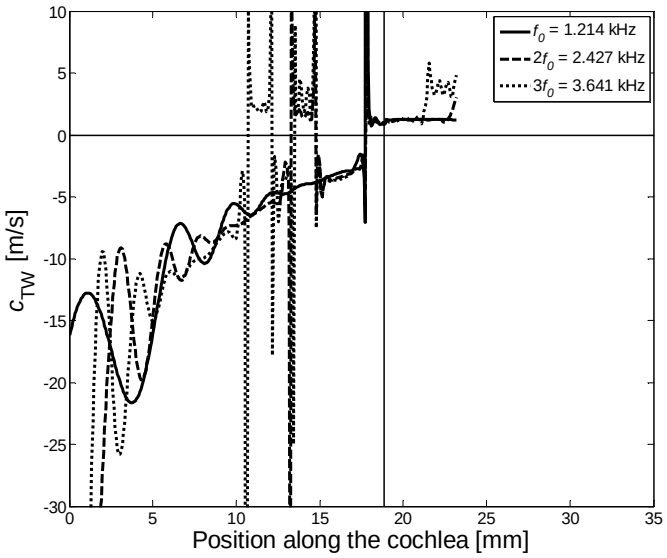


Figure 4.26: Steady state wave velocity of three limit cycle frequency components along the nonlinear BM. The solid line represents the linearly unstable frequency, whereas the dotted and dashed lines represent the $2f_0$ and $3f_0$ harmonics. Negative velocities represent backward TWs, while positive velocities represent forward TWs. A solid vertical line at 18.9 mm represents the location of the maximum BM response. The response is limited to the first 23 mm as the TWs become evanescent beyond this location.

A more rigorous approach to decomposing the BM response into forward- and backward-TWs would involve applying an analytical approximation, such as the WKB-method (e.g. Zweig, 1991; Neely and Allen, 2008). However, such methods were developed to approximate linear systems. This topic is given further consideration in Chapter 6.

The next few plots are presented to facilitate comprehension of the results given in Figure 4.17 and Figure 4.26, which showed the magnitudes, phases, and calculated wave velocities of several frequency components of $\dot{\xi}_b(x)$. A step-change in γ(x) resulted in a single linear instability, as described in given in Figure 4.12. Figure 4.27 shows only the magnitude and phase of the fundamental component of the BM velocity for clarity. Both the magnitude and the phase represent a superposition of forward- and backward- TWs. As described in equations (4.6) and (4.7), the slope of the TW phase with respect to position along the cochlea indicates its dominant direction of propagation. For instance, the wave is 'standing' at the top of the phase hill where $\partial\phi/\partial x = 0$ at x = 17.75 mm, shown by the dashed vertical line in Figure 4.27. The wave then propagates 'down' the hill away

in directions both basal and apical of this location. Thus, the backward TW dominates from the base of the cochlea to x = 17.75 mm.

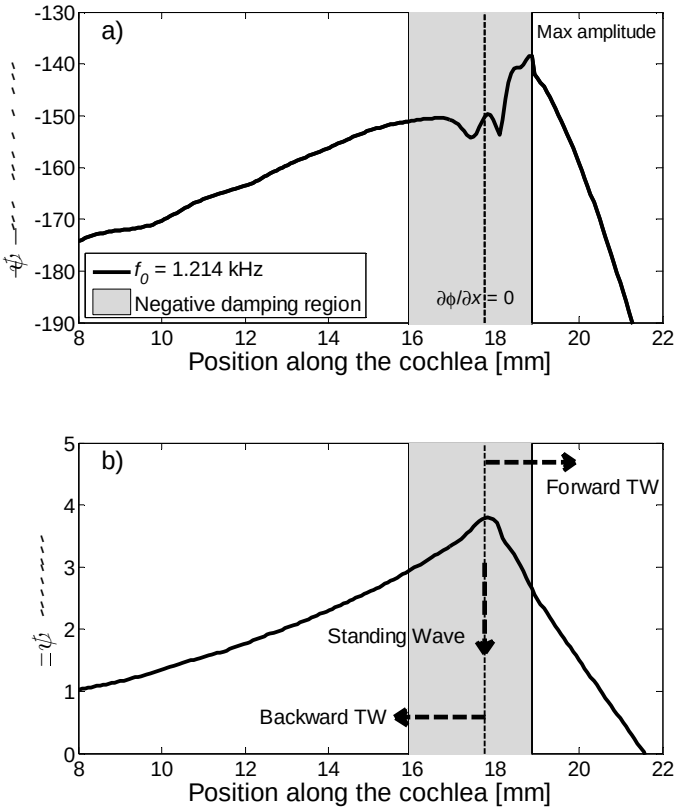


Figure 4.27.a-b: Steady state magnitude (a) and phase (b) of the BM response as a function of position at the unstable frequency. A dashed vertical line marks the $\partial\phi/\partial x = 0$ location at x = 17.75 mm, whereas a solid vertical line marks the location of maximum amplitude at x = 18.9 mm. The region of negative damping, given a stimulus frequency of 1.214 kHz in a baseline active cochlea, is shaded.

The wave speed is fastest near the base, as expected, and decreases at more apical locations. The propagating direction switches abruptly at x = 17.75 mm, after which it reaches a constant forward velocity of 1.2 m/s until it is extinguished, as shown in Figure 4.26. Note that the characteristic place of the BM response, shown by the solid vertical line in Figure 4.27, is apical of x = 17.75 mm. Furthermore, there is also a local maximum in the magnitude of each frequency component in Figure 4.17 just basal of x = 17.75 mm. Both of these peaks fall near the edges of the negative damping region, as shaded. These observations suggest that both forward- and backward-TWs are amplified in this scenario. In order to better inform this discussion, the annotations describing the direction of TW

propagation are overlaid upon a plot of the real part of the BM admittance in a baseline linear cochlea at the unstable frequency in Figure 4.28.

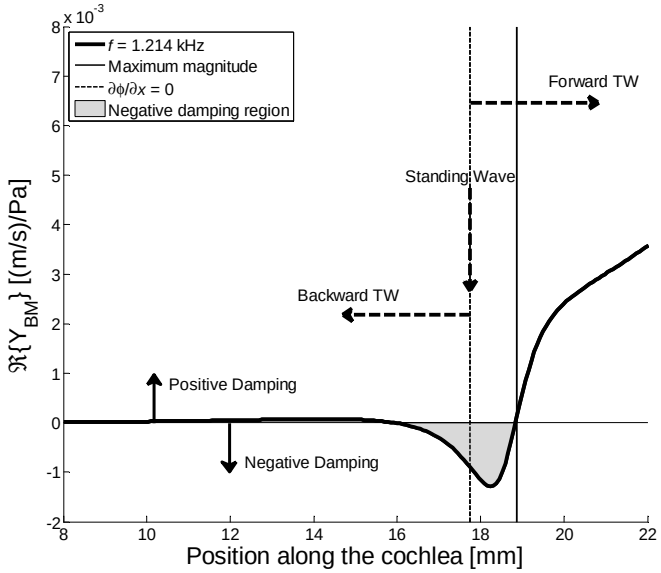


Figure 4.28: Real part of the BM admittance as a function of position along the cochlea, given f = 1.214 kHz. The peaks in both magnitude and phase of Figure 4.27 are plot here as solid and dashed vertical lines, respectively.

The forward TWs are amplified basal to the characteristic place, and peak in the region of the discontinuity. Backward TWs are then generated as forward-going wavelets reflect off of the perturbation; the backward TW is again amplified when heading toward the base which generates the second peak at $x \approx 16.5$ mm. Note that there is precedence for the amplification of backward TWs in a one-dimensional model, as demonstrated by Talmadge *et al*. (1998). However, the peak in the phase response represents the only position along the BM where the TW of the fundamental is 'standing,' as the amplitudes of forward- and backward- TWs are equal here. This results in a local peak in the magnitude at x = 17.75 mm. The local minima near this position are likely due to destructive interference of forward- and backward- TWs; these are only visible near the $\partial\phi/\partial x = 0$ place because the TW amplitudes are similar in this region.

It is interesting to note that Neely and Allen (2008) have reported a similar result in the magnitude and phase characteristics of a backward travelling pressure wave, as derived from a WKB approximation of wave propagation. They show that the backward TW is only supported when inhomogeneities (the stiffness as a function of position was

perturbed) are present along the BM. Furthermore, the peak in the phase response is basal of the peak in the magnitude response, just as in the simulations presented here. They conclude that 'the round trip delay of an SFOAE may be less than the twice forward delay to the characteristic place,' based upon the phase plot (Neely and Allen, 2008). It is, however, difficult to envision this given the normal propagation of slow waves along the BM.

A more likely interpretation is the one given above—the phase plot represents the combined effect of forward and backward components; at a given location, the positive or negative slope of the phase results from the dominant component. In Figure 4.27, the region between $\partial\phi/\partial x = 0$ and the maximum magnitude is dominated by a forward-amplified wave which shifts the phase response because $\partial\phi/\partial x$ is now negative. Thus, the round-trip delay is still twice the forward delay, but this is no longer apparent from the phase plot as the backward TWs are hiding 'under' the forward TWs in the overlap region basal to the characteristic place. This is illustrated in Figure 4.29, where the phases of three TWs are shown. The linear baseline cochlea only shows forward travel; its reflected response is plot by flipping its sign and DC-shifting it such that the backward TW phase intersects with the forward TW phase at the location of the discontinuity. When the unstable cochlea's phase is DC-shifted by -9 cycles (one complete forward-and-backward trip to the site of the discontinuity), there is good agreement between the backward TW's phase and the reflected wave's phase. Note the slight disagreement near the $\partial\phi/\partial x = 0$ location; this is believed to be due to the relatively similar amplitudes of the forward- and backward- TWs in this region, which may slightly affect the phase results.

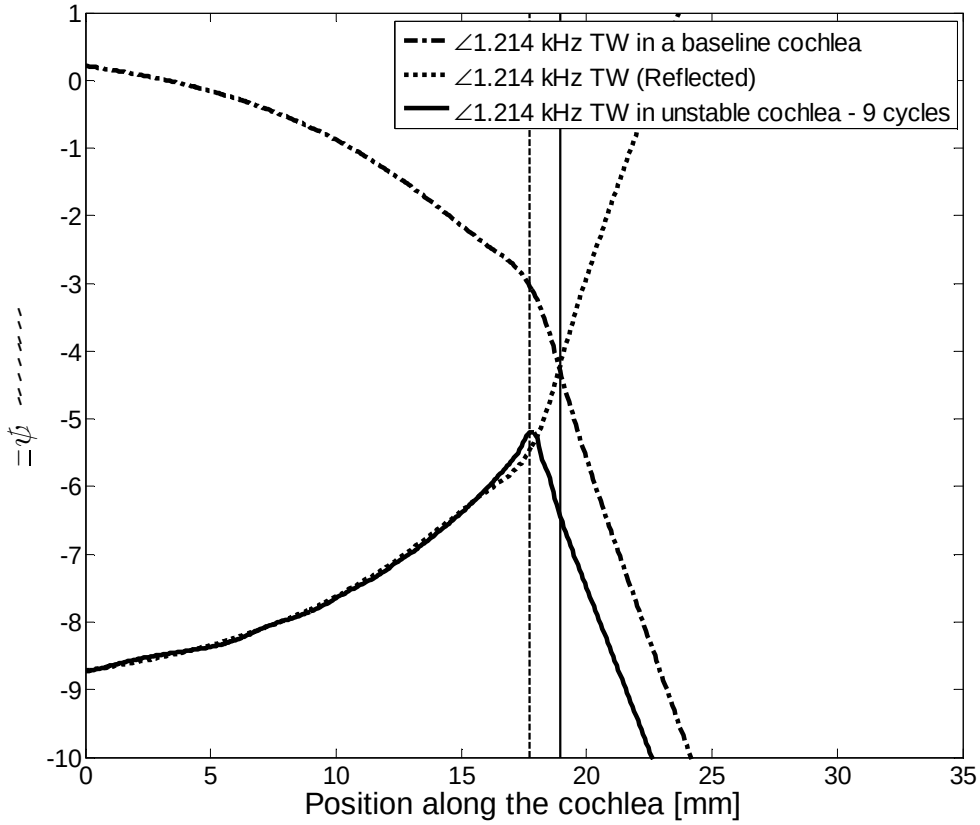


Figure 4.29.a-b: Phase of 1.214 kHz TW (dash-dot) and its reflected response at the location of the discontinuity (dotted). The steady state phase of the 1.214 kHz DFT in the unstable cochlea is shifted by -9 cycles and plot as a solid line. A dashed vertical line marks the $\partial\phi/\partial x = 0$ location; a solid vertical line marks the location of maximum amplitude.

The dynamics of the $2f_0$ and $3f_0$ harmonics, shown in Figure 4.30, are somewhat more difficult to interpret as the directions of the TWs appear to change multiple times. The fact that the dominant TWs of the harmonics switch directions at the same location basal of the strongest peak in the magnitude suggest that there is distortion reflected off of the discontinuity, which then becomes amplified at its best place. Thus, there appears to be a combination of both distorted amplification and amplified distortion taking place near the negatively damped region of the fundamental. A similar pattern of activity also exists at approximately 14 mm in Figure 4.30, near the characteristic place of $2f_0$.

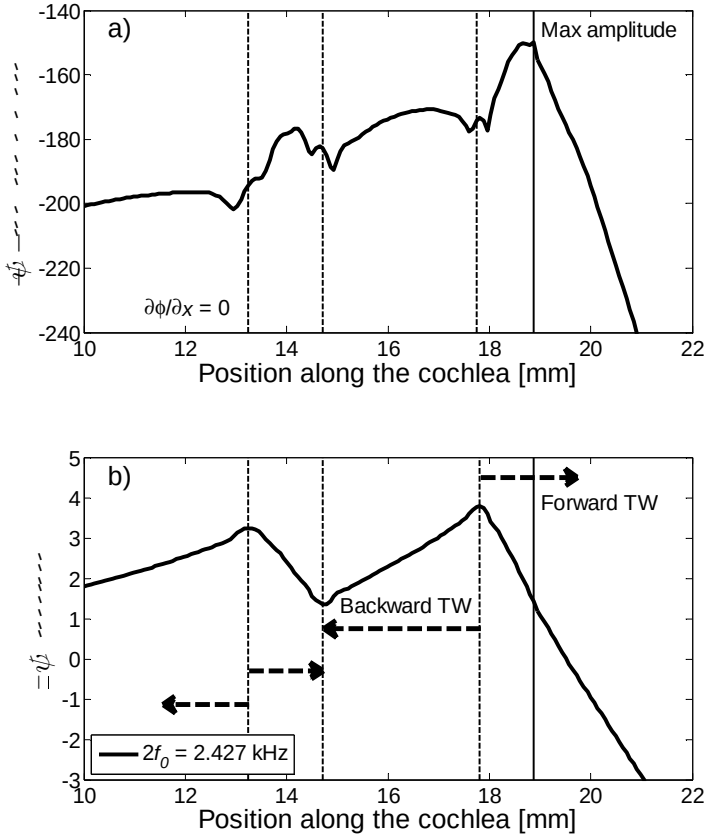


Figure 4.30.a-b: Steady state magnitude (a) and phase (b) of the $2f_0$ harmonic response of the unstable frequency.

Basal to x = 13.25 mm, the $2f_0$ component is dominated by a backward propagating wave. In the region where 13.25 < x < 14.7 mm, the TW is forward propagating. In this case, it is possible that $2f_0$ distortion generated near the characteristic place of the fundamental is propagating backward to the base and then reflecting forward to be amplified at the $2f_0$ characteristic place. However, it seems also plausible that the $2f_0$ distortion is being generated near its region of negative damping. This is analogous to the appearance of a 6 kHz ($2f_0$) distortion component near its characteristic place when 3 kHz sinusoidal stimulus is applied to a baseline model as in Figure 3.18. The $3f_0$ distortion component has a third phase 'hill' near its characteristic place; it is likely that the same mechanism responsible for the second 'hill' the $2f_0$ phase is at work there. The bottom line is that the exact nature of DP component generation, propagation and amplification in the cochlea remains an open area of research.

There are many unexplained features of nonlinear systems with multiple linear instabilities, such as that studied in Section 4.3.2. Is it possible, for instance, to predict

which linear instabilities will become limit cycle oscillations in the nonlinear simulation? Of the five linearly unstable poles generated by the windowed-perturbed $\gamma(x)$ given in Figure 4.19, only three persisted as strong limit cycles at the end of the simulation. The lowest frequency instability also had the smallest undamping ratio and yet was the second-strongest limit cycle. This is perhaps a physical manifestation of the phenomenon of upward spread masking, where lower tones effectively mask the detection of a higher-frequency tone. However, some adjacent linearly unstable poles resulted in limit cycles, while others did not. The real parts of the BM admittances of the three limit cycle frequencies in a baseline linear cochlea are superimposed on one plot in Figure 4.31.

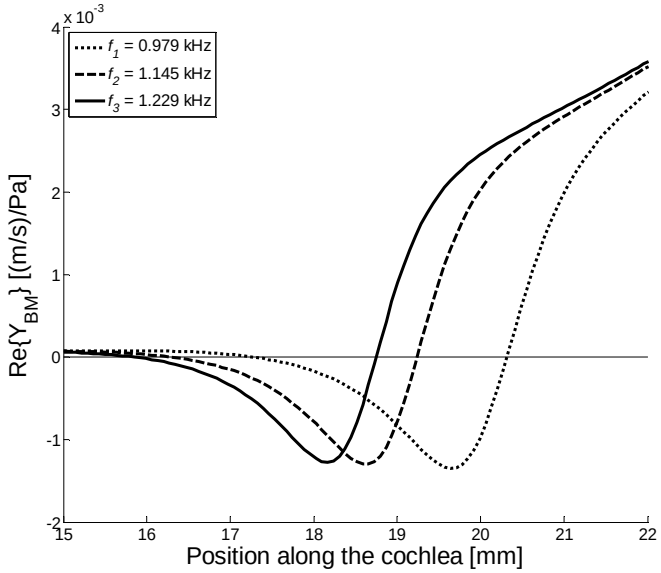


Figure 4.31: Real part of the admittance of three linearly unstable frequencies.

Figure 4.31 shows that there is significant overlap between the negative-damping regions of the f = 1.145 and 1.229 kHz tones, though neither tone suppressed the other. The two unstable poles at these frequencies had the highest undamping ratios of the instabilities, so perhaps this contributed to their persistence. It is likely that a combination of regions of amplification overlap, linear undamping ratios, and unstable frequency distributions may all contribute to or detract from the mutual suppression of adjacent unstable poles. One of the primary features of interest here is the spacings between limit cycle oscillations.

Figure 4.25 shows that the spacings between the selected limit cycles are amazingly consistent across frequency. Recall that the three linear instabilities that persist at steady

state are 0.166 and 0.084 kHz apart. Any nearby intermodulation distortion products will also be at integer multiples of those distances from the primaries. For instance, the nearest distortion product frequencies of limit cycle oscillations at 1.0 and 1.1 kHz would be at 0.9 and 1.2 kHz. This would result in Δf spacings that fall at regular intervals, similar to that shown in Figure 4.25.b. The validity of this line of reasoning is supported by reports in the literature of measured SOAEs that appear to be the result of cubic difference tones generated by 'primary' SOAEs (Burns *et al*., 1984; Whitehead, 1988).

Consider a large number of linearly unstable poles that arise given a 'dense' distribution of inhomogeneities. Even if the occasional linear instability is suppressed by a neighbour, a distortion product is likely to be generated nearby due to the next higher two (or previous lower two) instabilities in frequency. Thus, though linear reflection is no longer the mechanism giving rise to all limit cycle oscillations in the cochlea, the local PMD between their frequencies would still be expressed. This phenomenon nevertheless requires that a 'dense' distribution of inhomogeneities be present in the first place to fix the regular underlying spacings between linear instabilities, as dictated by λ_{peak} and the frequency-to-place map. Thus, it would appear that human spontaneous otoacoustic emissions are amplitude-stabilized cochlear standing waves and their intermodulation distortion products.

The only flaw in such an argument deals with the magnitudes of the distortion-generated limit cycle oscillations, which are approximately 40 dB below that of the primaries in the presented simulation. This would put the distortion components below the noise floor in all but the most sensitive recording regimes. Nevertheless, the relative amplitude of the two primary tones is known to have a strong impact upon the level of the distortion product (Hall, 2000). In addition, applying inhomogeneities to the entirety of the model may more strongly reflect the distortion components from different positions in the cochlea. As such, there may be other factors contributing to the overall response that have not yet been considered; further simulations may show distortion-generated limit cycle oscillations of higher magnitudes.

It should be noted that the pressure data presented in this chapter was calculated only at the base of the model. It is possible to determine the limit cycle amplitude in the ear canal by passing this signal through the two-port network, as described in Appendix A. However, so little is understood about the precise generation mechanisms in the cochlea

that this additional linear transform is left to future work. However, it is known that the middle ear primarily acts as a band-pass filter, with a pass region of approximately 0.5 – 2 kHz. At the peak near 1 kHz, the reverse pressure transfer function is approximately -30 dB SPL in magnitude. The stronger limit cycles generated in this chapter's simulations had magnitudes at approximately 0 dB SPL at the base, which are clearly too low to be measured in the ear canal. However, it is very simple to adjust the magnitudes of limit cycles by revising the overall saturation point, δ, in the Boltzmann nonlinearity. A higher overall saturation point will simply shift up the (steady state) saturated response accordingly. Furthermore, results in the previous chapter also suggest that this value (currently 1 nm) is too low.

There is a great deal of future research that can be pursued following these results; these ideas are discussed in Chapter 6. The next chapter takes a step back from the boundary of stability to analyze the system's response to clicks. Both linear and nonlinear simulations are performed, but all cochleae are stable to simplify the interpretation of results.

Chapter 5

Click Evoked Otoacoustic Emissions

Along with spontaneous emissions, the existence of physiologically vulnerable and level-dependent transient evoked otoacoustic emissions (TEOAEs) supports the notion that a nonlinear, active mechanism is at work in the human cochlea. To be consistent with the rest of the thesis, a TEOAE is defined here as a variation in pressure at the stapes or in the ear canal that is generated in the cochlea.

As discussed previously, the mammalian cochlea distributes and amplifies the energy of a signal's frequency components to various positions along the CP, according to its frequency-to-place map. It is also understood that deviations from the smooth spatial variation of the CA, or other mechanical properties of the CP, can cause reflections of forward-travelling waves in the cochlea. Thus, when an impulsive excitation is introduced at the base of the cochlea, the various frequency components of the click reach maxima at their characteristic places and are most strongly reflected from these locations. However, the group delay, or the time required for a TW to reach its characteristic place, is frequency-dependent. The round-trip travel time for a wavelet to propagate from the base to its characteristic place and back is thus referred to as the frequency-dependent TEOAE latency. The highest frequency components return to the stapes first because they peak sooner in time and closer in longitudinal distance from the base, relative to the lower frequencies. This finding was first presented by Kemp (1978), who recorded TEOAEs in humans following both clicks and four cycle-long tone-burst stimuli from the ear canal.

Numerous investigations have since verified these results (Hall, 2000; Sisto and Moleti, 2002).

There are a number of clinical methods of measuring TEOAEs. These are broadly classed as derived nonlinear (DNL) and linear approaches, both of which seek to eliminate the linear stimulus and associated artefacts while preserving the nonlinear OAE. In the DNL paradigm, the difference is taken between the recorded waveforms of two (or more) clicks of varying stimulus amplitudes, typically 6 dB or 9 dB apart. The response due to the lower amplitude click is scaled up by the corresponding increase in stimulus level for the higher amplitude click; the difference is then taken between these two quantities such that the linear components of each are eliminated. These clicks are typically 100 μs in duration and are repeated every 12.5 to 30 ms. As the biological CA behaves linearly when tonally excited up to approximately 30dB SPL, any attempt to resolve a click-evoked otoacoustic emission (CEOAE) using nonlinear cancellation must operate at a higher level, otherwise both the stimulus artefacts and the CEOAE will be lost when the difference of the signals is calculated. For example, the UK neonatal screening programme recommends a stimulus level of 80-88 dB peak-equivalent SPL (Hall, 2000). The results of many such periodic clicks are averaged in time in order to reject noise.

The linear paradigm of CEOAE measurement also reduces noise by time-averaging the signal recorded in the ear canal. In order to remove the stimulus artefact in the linear paradigm, the difference between its short latency and the CEOAE's longer latency is exploited. The artefact, which is due to ringing in the ear canal and the middle ear, typically decays to negligible levels relative to the CEOAE within 3 to 8 ms following the onset of the click. Thus, up to the first 5 ms of linear time-averaged TEOAE measurements are windowed out to remove the artefacts (e.g. Tognola *et al.*, 1999). Both linear and nonlinear approaches have advantages and drawbacks.

DNL methods greatly reduce the chance of mistaking a click artefact for a genuine OAE. In addition, nonlinear methods are able to resolve the highest frequency components of CEOAEs which occur first in time following the stimulus. In contrast, linear methods window out this section of the signal to eliminate stimulus contamination; linear methods can, however, maintain a higher signal-to-noise ratio (How and Lutman, 2007). Furthermore, at amplitudes where the CA is partially- or fully-saturated, the cochlea behaves differently from its low-level, linear response. One would expect to see the

greatest ratio of CEOAE level to stimulus amplitude when the CA is linear, though the noise due to physiological and environmental sources precludes these low-level measurements in practice. One effect of increasing stimulus amplitudes is that the TEOAE latency begins to decrease (e.g. Tognola et al, 1999). This shift toward earlier latencies has also been observed in the envelope of click-evoked BM motion *in vivo* (e.g. Recio *et al.*, 1998).

Thankfully for the cochlear modeller, the myriad of practical and philosophical difficulties associated with clinical measurement and analysis can often be sidestepped. The only noise associated with mathematical simulations of the cochlea is due to errors generated during the solving of the system's ordinary differential equations. This is minimized by setting the relative and absolute error tolerances of the ode solver to low values, with 10^{-8} and 10^{-17}, respectively being used below. All of the traditional nonlinear and linear methods of TEOAE isolation can be applied in post-processing of simulation results, in addition to some which would be impossible in a clinical study. The component of the pressure at the base that is due to the stimulus can be isolated by exciting a baseline cochlear model, a model with no inhomogeneities, with the same input. This is then linearly subtracted from the output of a model with inhomogeneities, thus leaving the CEOAE. This is defined as a 'directly-determined' (DD) CEOAE; this method can be applied in both linear and nonlinear simulations.

The stimulus used for the following experiments is a rectangular pulse of volume displacement simulated in the ear canal and lasting 100 μs. The output acceleration at the stapes that is generated by a pulse in volume displacement of $1*10^{-16}$ m^3 in the ear canal is approximately equivalent to an RMS magnitude of output acceleration at the ear drum generated by a 0 dB peak SPL excitation lasting 100 μs. Further details on the generation of this stimulus and can be found in Appendix A.

This chapter is divided into results from linear and nonlinear simulations. The DD method is applied to isolated CEOAEs in linear simulations, and both DD- and DNL-CEOAEs are calculated given nonlinear simulations. The linear section presents click responses for the baseline model and a large set of inhomogeneous but stable cochleae. The CEOAE latencies of the perturbed models are analysed and compared with predicted latencies. The nonlinear section presents click responses of increasing amplitude for a baseline model, and a model with the same perturbation in the feedback gain as used in the

linear analysis. Finally, models with perturbations in the scaling of the nonlinearity in the CA as a function of position are also analysed for comparison. The latencies of CEOAEs generated from models with perturbations are then analysed and discussed.

5.1 Linear Simulations

Linear simulations of clicks are informative in a number of ways. The response of a healthy mammalian cochlea can be considered linear at low levels, less than approximately 30 dB SPL, and again at high levels above approximately 100 dB SPL, (Robles and Ruggero, 2001). Thus, an active linear model responds similarly to the cochlea at low levels, whereas a passive linear model responds similarly to the cochlea at high levels. However, it is important to note that these similarities should not be taken to suggest that a linear model can encompass all of the properties of a nonlinear cochlea, even given these restrictions. For instance, at high levels of stimulation, it is possible that the gain provided by the OHCs is saturating in only one portion of the cochlea, or for only a short period of time.

Nevertheless, many characteristics of the cochlea can be captured with a linear model, hence the persistence of such formulations in the literature (de Boer, 1996). One such feature is the variation of SFOAE fine structure as measured in the ear canal within small ranges of frequency, which is only visible at low levels (e.g. Zwicker, 1990). It is within this low-level, linear range that the greatest amount of reflection per unit input is generated. Unfortunately, the click-reflected OAE is exceedingly quiet at low levels where the cochlea is linear, and often falls below noise thresholds when measured clinically. The cochlear model, in contrast, generates very clean results with much higher signal-to-noise ratios. The results presented in this subsection were all generated by a standard 100 μs click in volume displacement presented in the ear canal.

5.1.1 BM Responses

Figure 5.1(a-d) shows mesh plots of BM velocity as a function of time and position in the cochlea. The (a) and (c) panels represent baseline active and passive cochleae, respectively. Panel (b) shows the response of a stable active model with dense, 0.75% peak-to-peak inhomogeneities in $\gamma(x)$. The stability of this model is shown in Figure 5.2.

Figure 5.1.d shows the difference of the random response (b) and the baseline response (a) in the cochlea; only motion due to reflections remains.

Backward travelling waves are visible in Figure 5.1.d for t and $x > \tau_{\mathrm{TW}}(x,t)\big|_{\gamma=1}$, plotted as a solid line in this panel. This is the group delay, or predicted latency to each characteristic place across frequency, as presented in Chapter 2,

$$\tau_{\mathrm{TW}}\left(x_{cf},\omega_{cf}\right)=\int_0^{x_{cf}}\frac{\partial\,\mathrm{Re}\left(\kappa_{\mathrm{TW}}\right)}{\partial\omega}dx\,. \tag{2.37}$$

Note that the group delay at a given position is longer for the active cochlea, as shown in Figure 2.21. After a given wavelet peaks at its characteristic place (somewhere along the black line, depending on its frequency), it is reflected back toward the stapes and reaches the stapes at approximately twice its forward delay, $2\tau(x,t)\big|_{\gamma=1}$. When it reaches the stapes, the impedance mismatch between the cochlea and the middle ear causes a portion of the wavelet's energy to be reflected back toward its characteristic place. A dash-dotted black line shows $3\tau(x,t)\big|_{\gamma=1}$. Forward travelling waves are discernable along the three times group delay line. The wavelet, having now been reflected twice, is again amplified as it approaches its characteristic place for the second time, along the dash-dotted $3\tau(x,t)\big|_{\gamma=1}$ curve.

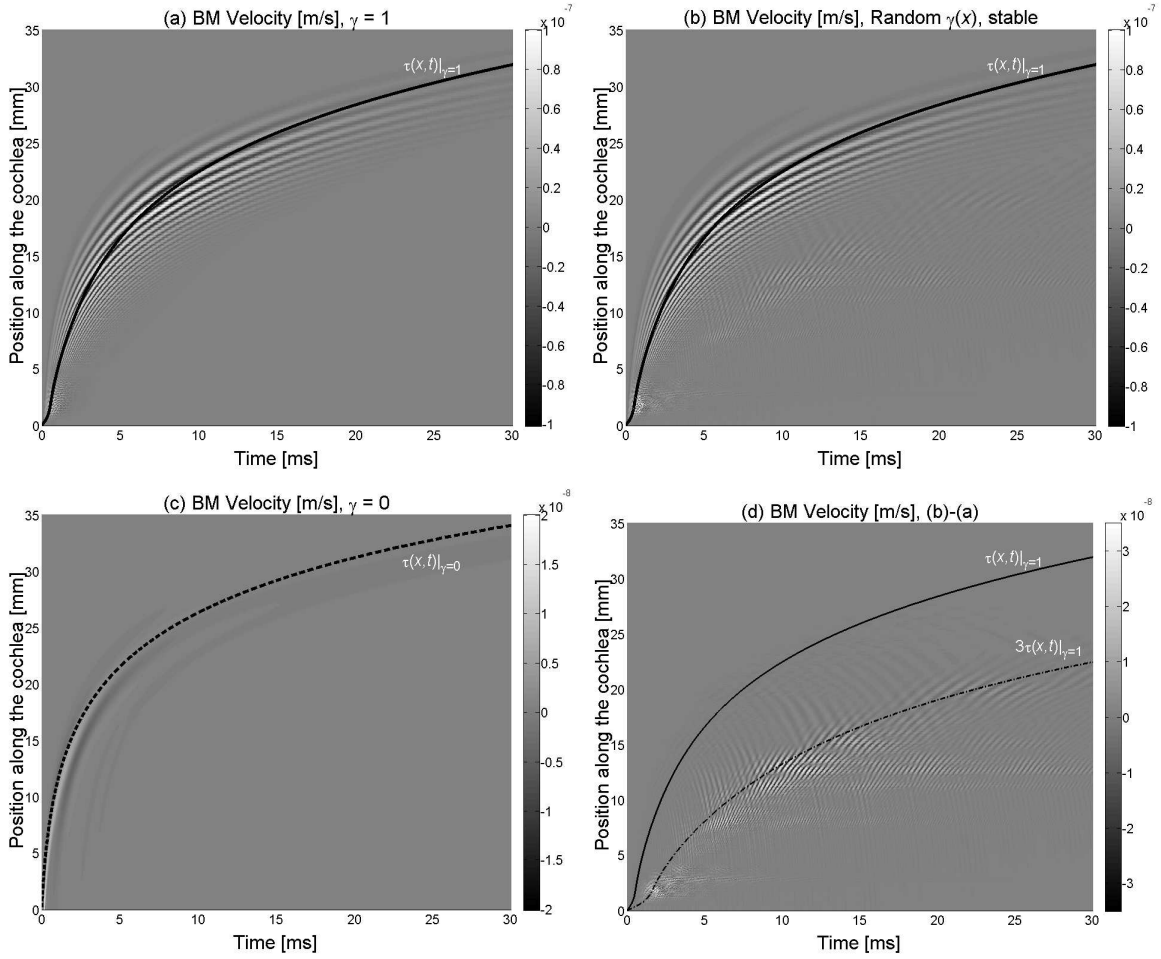


Figure 5.1.a-d: Response of linear cochlear models due to a 100 μs pulse of volume displacement equivalent to 20 dB SPL: a) baseline active model; b) active model with inhomogeneities; c) baseline passive model; d) difference between (b) and (a). The predicted group delay for an active baseline cochlea are overlaid in solid black for (a,b,d). The predicted group delay for a passive baseline cochlea is overlaid in dashed black for (c). Three times the group delay of the active case is overlaid in dash-dotted black in (d).

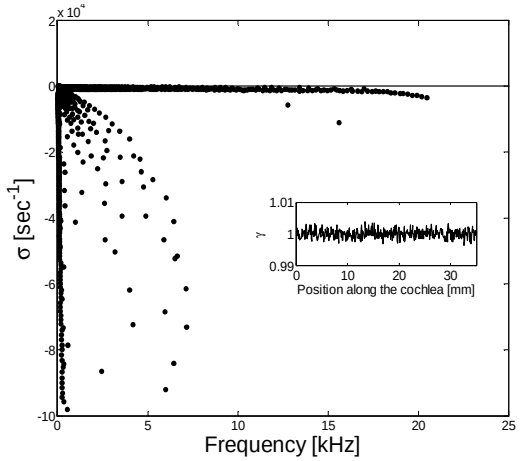


Figure 5.2: Stability plot for a cochlear model with dense, random 0.75% peak-to-peak inhomogeneities in $\gamma(x)$. This system is stable, and its linear click response is shown in Figure 5.1.a-d.

In anticipation of the nonlinear responses presented in the following subsection, Figure 5.3.a-d shows the displacement at a single position along the BM when excited by a pulse in volume displacement in the ear canal, given uniform variations in $\gamma(x)$. The x-scale is presented not in terms of time, but rather time scaled by the characteristic frequency (cycles) at that location, as by Shera (2001b). This is done both to qualitatively discern the amount of scaling symmetry that exists in the model and also to facilitate comparison between the two locations. The y-scale of the data is normalised by maximum displacement at each γ to enhance the visibility of the variations in the waveform.

There are two particularly notable differences between (a) and (b): the decay rate of the envelope at higher gains, and the shape of the waveform in the passive cases. The slower rate of decay in the envelope in (a) vs. (b) is to be expected as (a) is more basal, and thus more sharply tuned. The differences in the shapes of the passive waveforms, and thus the underlying slow responses of the active waveforms, are somewhat surprising. The first positive peak and second negative peak in the passive response of (b) is much more compressed in normalised time than those of (a). This suggests that basal points in this model are relatively more heavily damped (for the passive case) than those in the middle region of the model.

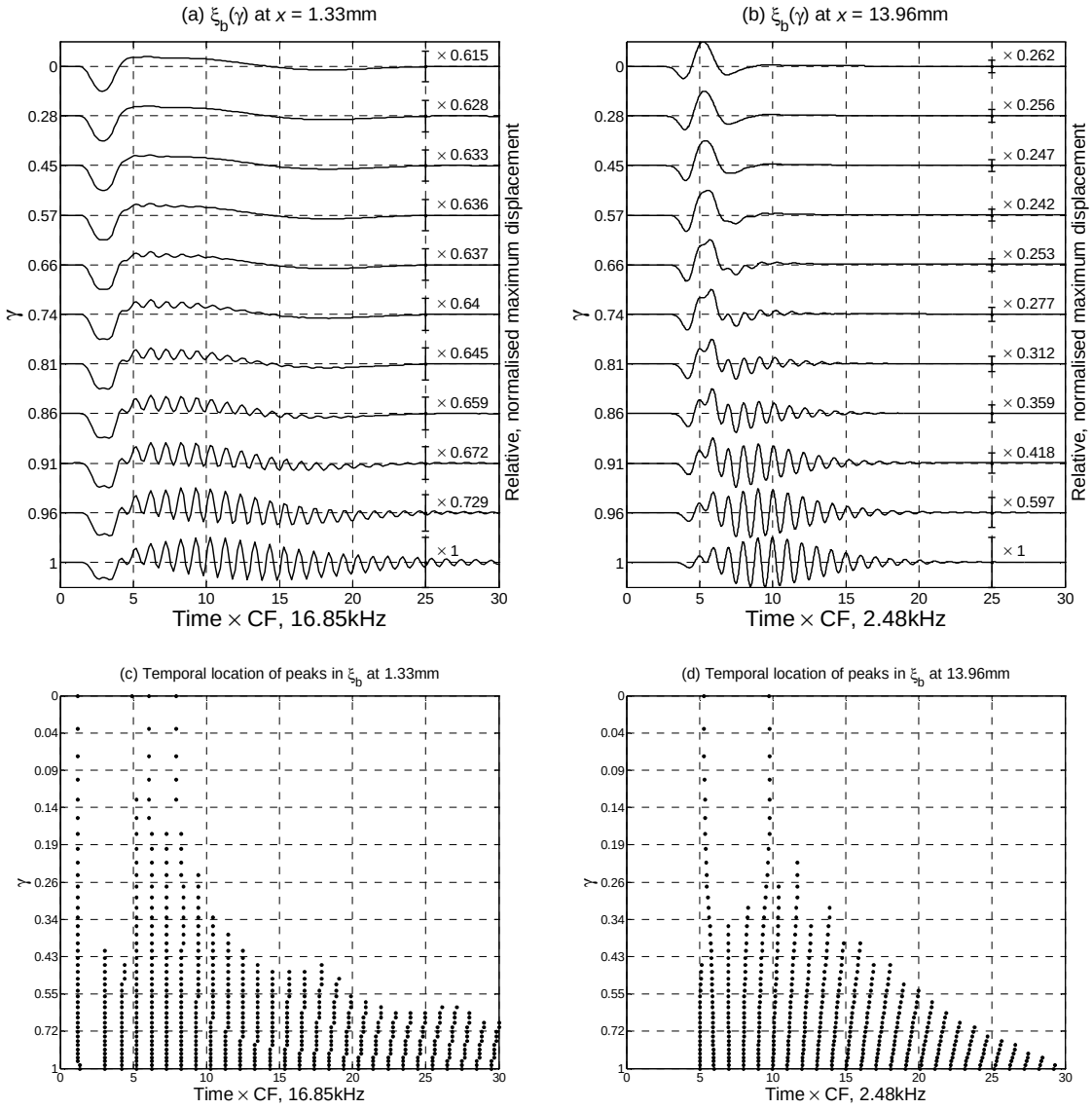


Figure 5.3.a-d: (a) and (b) show normalised BM displacement at 1.33 mm and 13.96 mm in a baseline linear cochlea due to a standard click with variations in gain from fully active (γ = 1, bottom) to entirely passive (γ = 0, top). The relative scales of the maximum displacements are plotted as error bars and scalar factors on the right. (c) and (d) plot the locations of the peaks of the response given the same progression from active to passive.

Figure 5.3.a-d shows the locations of the positive peaks in the displacement waveform as $\gamma(x)$ is varied. According to measurements made in the cochleae of live animals (e.g. Recio and Rhode, 2000), there is a near-invariance in the temporal fine structure of BM responses with varying stimulus level. Simply put, the positions of the peaks in the BM motion are level-independent, though the peak of the envelope shifts to earlier times with increasing intensity. Shera (2001) draws the conclusion that this result suggests OHCs are tuned to the same frequency as the natural response of the passive cochlea at any given point. The Neely and Kim (1986) formulation falls into the category of cochlear models that Shera denotes '*M*<,' where the poles of the micromechanical model move in frequency with changes in the feedback gain. Figure 8 of Shera (2001) illustrates how the peaks of an impulse response move when the active gain of such a model is varied. The same trend of increasing latencies with decreasing gains is observed in Figure 5.3.a-d.

5.1.2 CEOAEs

In this subsection, click evoked emissions are simulated. The pressure at the base of a linear perturbed cochlear model is calculated following a click stimulus. The response at the base is then transformed to obtain the response in the ear canal.

Figure 5.4 shows both short and long latency windows of the pressure at the stapes and in the ear canal given (I) a baseline model, and (II) the perturbed model of Figure 5.2. The (a) and (b) panels show the pressure at the stapes and ear canal for the first 5 ms of the response, whereas the (c) and (d) panels show the same data but for 3 ms < t < 30 ms. In most methods of linear OAE extraction, the waveform is windowed as in (d), to remove the response due to the stimulus (b) which decays away within several milliseconds. The short latency plots (b) are qualitatively similar to clinical measurements made by Harte (2004). There is almost no discernable difference between the short latency plots I(a-b) and II(a-b), as the stimulus itself is so dominant. However, whereas the long latency response of the baseline cochlea, I(c-d), decays away to imperceptible levels within the given time frame, the long latency response of the perturbed model, II(c-d), shows a great deal of activity due cochlear reflections.

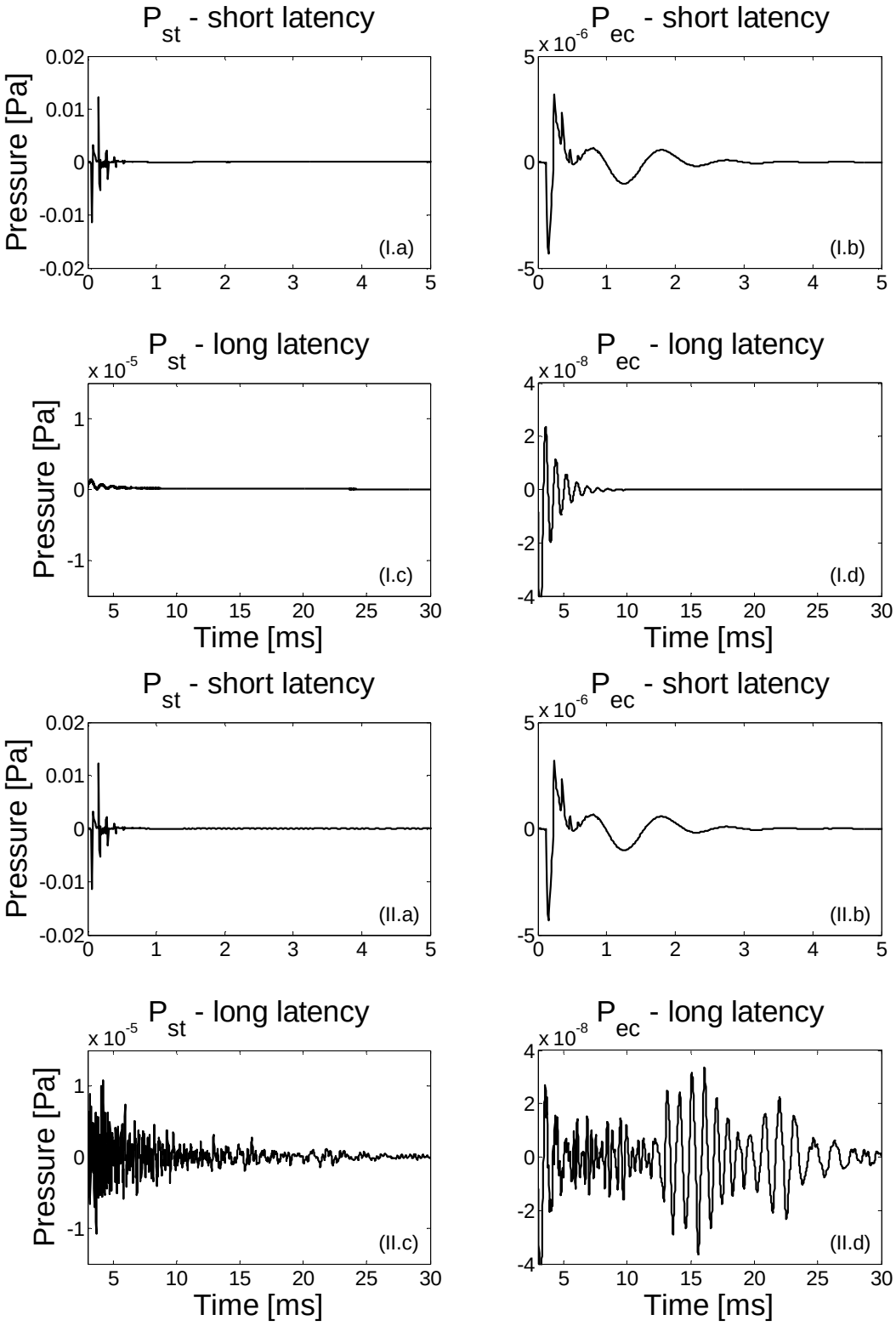


Figure 5.4.I-II.a-d: (I) The response of the baseline active cochlea given a standard click at 20 dB SPL, compared with (II) the perturbed, stable model of Figure 5.2. Pressure at the stapes (left panels) and the ear canal (right panels) given short time windows (a and b) and long time windows (c and d).

Visual inspection of Figure 5.4.II.d suggests that the instantaneous frequency of the CEOAE is a decreasing function of time. Although the cochlear model studied in this thesis is neither scaling symmetric, nor is its fine structure response invariant with changes in γ, the frequency-to-place map of the active baseline model is quite similar to that of Greenwood (1990) over much of the cochlea. Thus, it is expected that reflections due to inhomogeneities in the cochlear impedance as a function of position would have dominant latencies that are twice as long as the group delays of those frequencies in the cochlea, when measured at the stapes. Thus, the first (and presumably strongest) TEOAE latency should be $2*\tau_{TW}$ plus any time delay due to forward/reverse middle ear transmission.

It would be interesting to analyse the time-variation of the frequency spectrum of the TEOAEs using a short-time Fourier transform. Unfortunately, the limited window time of decaying CEOAE activity restricts the resolution in frequency such that a spectrogram generated by this method does not provide much useful information in practice. The latency of various frequency bands is, nevertheless, a commonly measured feature of CEOAEs. In order to determine the OAE latency as a function of frequency, the wavelet transform is applied here using MATLAB's continuous wavelet transform (CWT) function, `cwt`.

Wavelet analysis is a commonly used method to determine OAE latencies (e.g. Sisto and Moleti, 2007). This is because wavelet analysis is able to better resolve the time at which certain frequencies are expressed in a waveform, compared to traditional short-time Fourier analysis. The wavelet transform can be considered as a convolution of a signal with a 'mother wavelet' which is scaled in time to produce varying centre frequencies,

$$CWT(a,b) = \int_R s(t) \frac{1}{\sqrt{a}} \Psi\left(\frac{t-b}{a}\right) dt \quad , \tag{5.1}$$

where $s(t)$ is the signal being decomposed, a is the scale (analogous to frequency), b is the position (in time), t is time, and ψ is the mother wavelet. A variety of mother wavelets have been successfully applied in OAE analysis (e.g. Wit *et al*, 1994). Empirical testing has shown that the choice of mother wavelet does not strongly impact the final results; the Morlet wavelet, a Gaussian-scaled sinusoid, is chosen here,

$$\Psi_{morlet}(x) = e^{-x^2/2} \cos(5x) \quad , \tag{5.2}$$

illustrated in Figure 5.5.

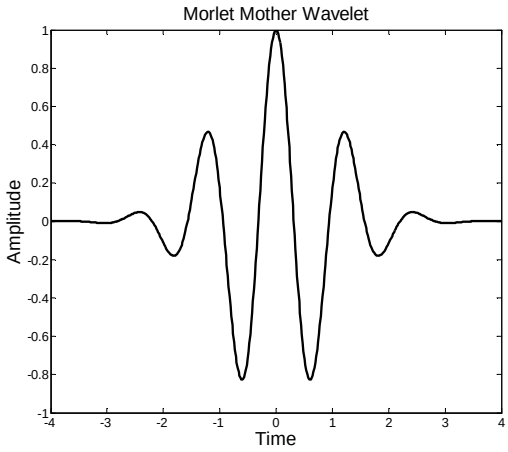


Figure 5.5: The Morlet mother wavelet.

The results of the wavelet analysis, applied to the OAE data presented in Figure 5.4.I-II.a-d, are shown in Figure 5.6.c-d. Latencies are plot along the horizontal axis while centre-frequencies are set decreasing along the vertical axis in order; this unorthodox format was chosen to emphasise the connection between the CEOAE latency and the cochlear group delay. The stimulus artefact was first removed by applying the DD-method of subtracting the stimulus response in the baseline model:

$$\begin{aligned} DD\ CEOAE_{st} &= p_{perturbed}\Big|_{st} - p_{baseline}\Big|_{st} \\ DD\ CEOAE_{ec} &= p_{perturbed}\Big|_{ec} - p_{baseline}\Big|_{ec} \end{aligned} . \tag{5.3}$$

The complete, DD CEOAE at the stapes and in the ear canal is shown in Figure 5.6.a-b. The *CWT* coefficients were computed as a function of time in 11 half-octave bands, extending from 0.5 kHz to 16 kHz. The time at which the maximum absolute value of the band-averaged CWT coefficient occurred was taken to be the latency in that frequency range. The predicted latency and the band-averaged latencies agree very well for this case, within 2% of all predictions.

The CEOAEs of 100 linear cochlear models with unique, random inhomogeneities were then computed. All of the inhomogeneities applied are 'dense,' and have high-wavenumber cut-off wavelengths varying from approximately 0.2 mm to 0.15 mm. A small peak-to-peak variation in gain of 0.75% was applied, and all distributions of gain were tested for stability before the simulations were run (unstable distributions were rejected). Although there was greater variability in the calculated band-averaged latencies, the results again matched up quite well with predictions, as shown in Figure 5.7. The

average latency in a given band never deviated more than 5% from the predicted latency at a particular frequency, and the predictions fell within ± one standard deviation of the averages at all but one band for each set of results.

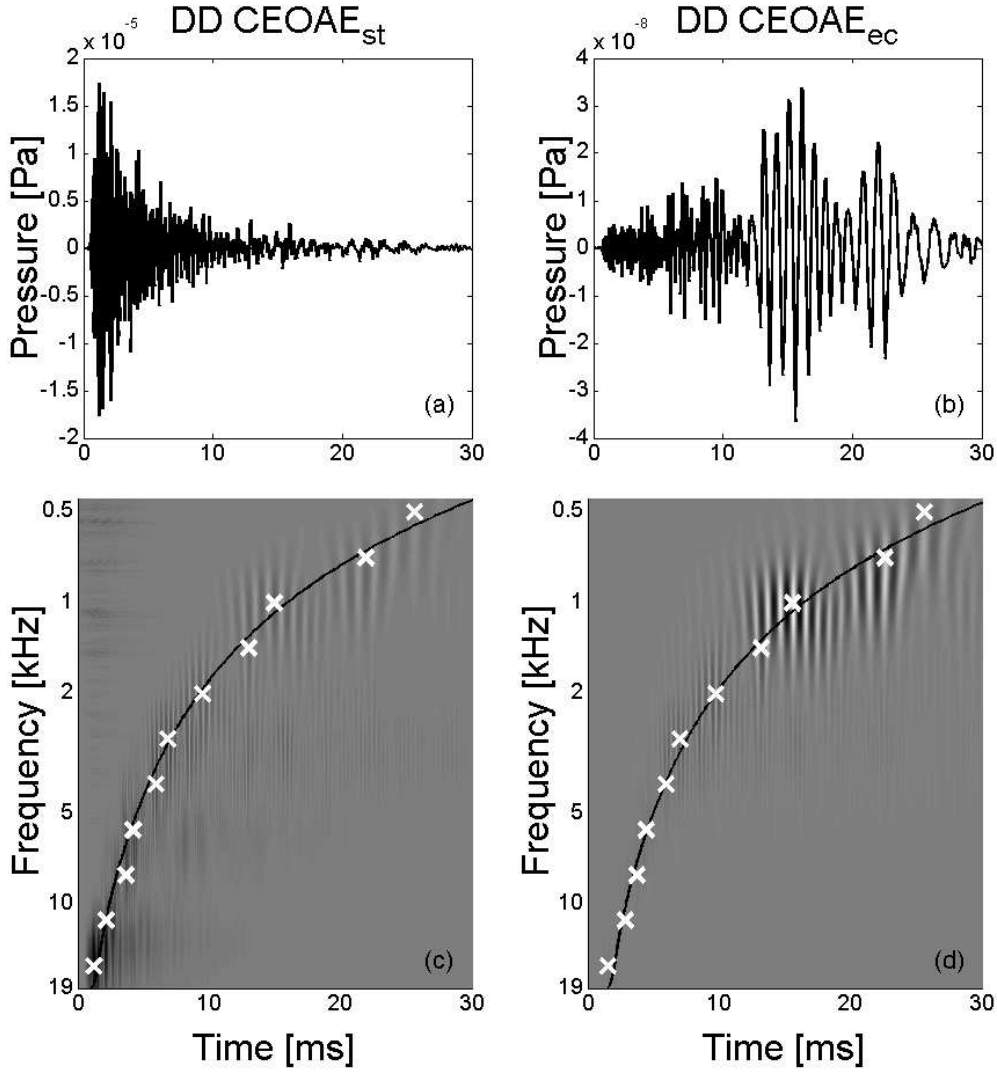


Figure 5.6.a-d: Directly determined CEOAE waveform at the stapes (a) and in the ear canal (b). Panels (c) and (d) show the continuous wavelet transform coefficients of (a) and (b), with the predicted latency overlaid in black (—) and band-averaged cwt maxima as white (×).

The results for the lowest band, 0.5 kHz, are somewhat skewed to earlier times due to the simulation time window; a longer simulation time should resolve this band more accurately.

When compared to clinical measurements of CEOAE latencies (e.g. Moleti *et al*, 2005), the average results from the linear model overestimate the latency in every frequency band. This is to be expected, however. Numerous studies have demonstrated

that there is a statistically significant correlation between CEOAE latency and stimulus level. This is well documented by Sisto and Moleti (2007). As the stimulus level is increased, the latency decreases more or less monotonically in each frequency band. As CEOAE measurements are typically made in a strongly saturating amplitude range, it is no surprise that measured latencies are shorter than those observed in the linear model. This trend is predicted by the revised Neely and Kim model, which shows monotonically decreasing $\tau_{TW}(\omega)$ with decreasing gain as shown in Figure 2.21.

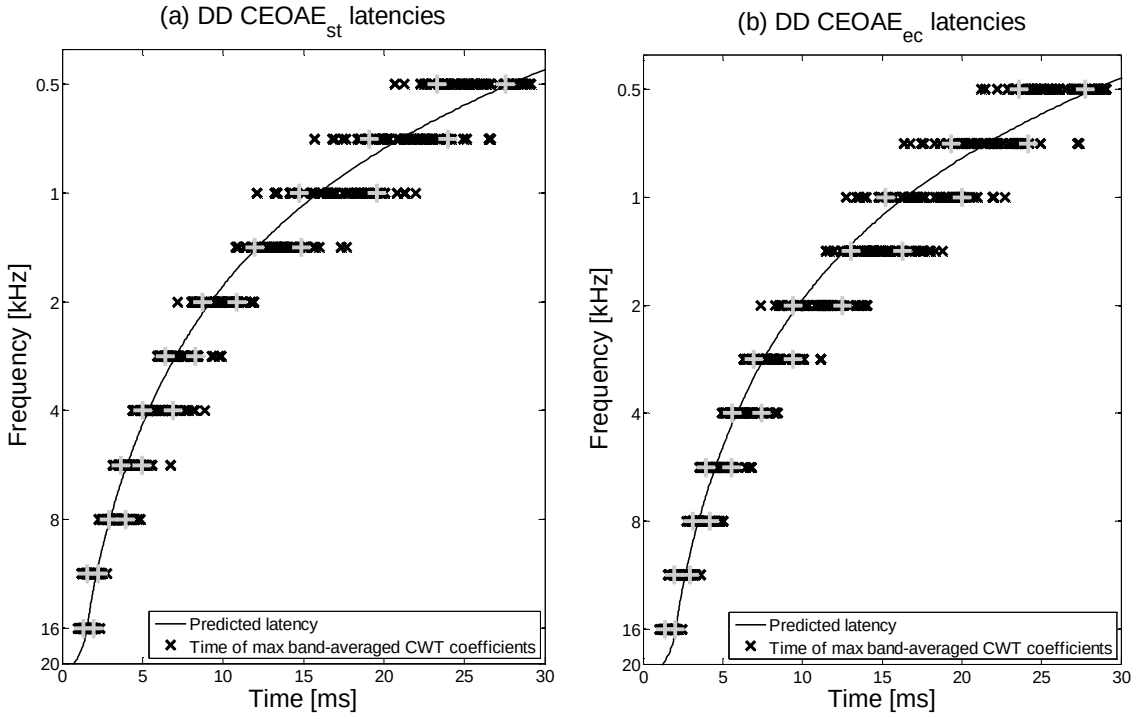


Figure 5.7.a-b: Predicted CEOAE latencies for an active baseline cochlea (solid line) compared to calculated, band-averaged latencies from simulated DD CEOAEs ('**x**'s) in perturbed cochleae. The response is compared at the stapes (a) and in the ear canal (b). Gray (+) symbols represent ±1 standard deviation about the mean. This figure represents the collected results of 100 linear models with unique, randomly varying $\gamma(x)$.

5.2 Nonlinear Simulations

The nonlinearity of the cochlea is one of its most remarkable features because this allows mammals to perceive a much wider range of loudnesses than would otherwise be possible. This subsection first presents the nonlinear model's BM responses to clicks, which are analogous to the linear results shown in Section 5.1. The CEOAE of a single perturbed model (the linear stability of which was presented in Figure 5.2) is then examined as the click amplitude is increased. At the end of this subsection, a simulation is carried out by locating the inhomogeneities not in the micromechanical gain as a function of position, $\gamma(x)$, but in the nonlinear saturation point of the active elements as a function of position, $\delta(x)$. This shows how reflections may arise due to non-smoothly varying parameters in the nonlinearity, as well as linear quantities.

5.2.1 BM Responses

Figure 5.8 shows mesh plots for nonlinear time domain simulations for clicks of increasing amplitudes applied to the perturbed model presented in Figure 5.2. At 0 dB SPL, the response of Figure 5.8.a-d is identical to that of Figure 5.1, the linear case. As the stimulus amplitude increases, the feedback force begins to saturate. Consequently, the CA's relative contribution to the motion of the BM is suppressed; this is comparable (but not exactly equivalent) to reducing the feedback gain in the linear model. At the highest stimulus levels, as in Figure 5.8, only the passive response is visible near the peak of the TW; this is qualitatively similar to Figure 5.1.c.

The same trend of motion resembling active responses shifting toward passive responses with increasing stimulus level is visible locally as well as globally. Figure 5.9 shows this pattern in the displacement at two locations along the BM. Again, as in Figure 5.3.a-d, the x-axis is plot in terms of time × CF (cycles) for easy cross-comparison. The variation of the BM displacement waveform in (a) and (b) is quite similar for the analogous linear simulation, shown in Figure 5.3.a-b. There are, however, subtle differences.

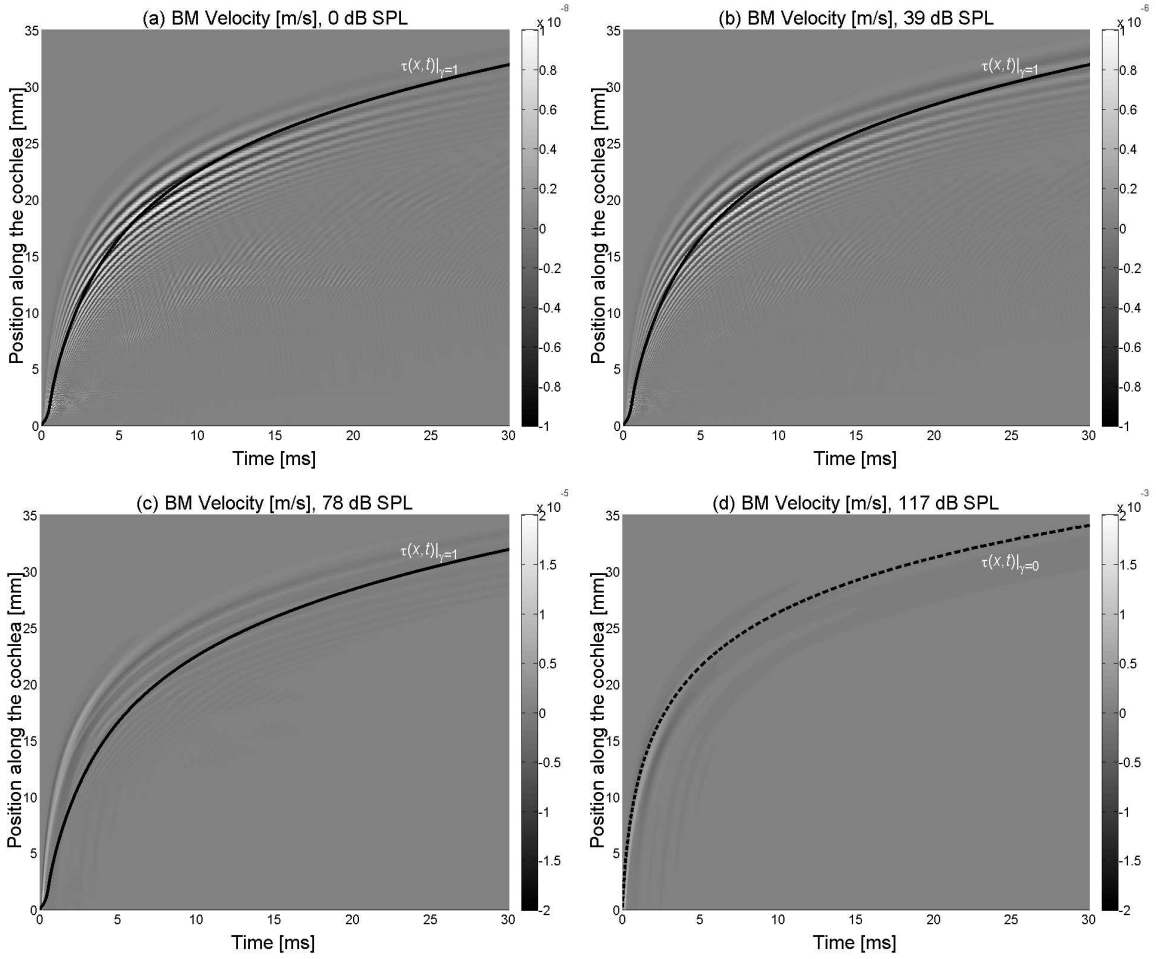


Figure 5.8.a-d: Nonlinear cochlear response of model presented in Figure 5.2 due to a standard click at 4 stimulus levels: (a) 0 dB SPL, (b) 39 dB SPL, (c) 78 dB SPL and (d) 117 dB SPL. The cochlear group delay is plotted for a baseline active cochlea as a solid line (a-c). The cochlear group delay is plotted for a baseline passive cochlea as a dashed line in (d).

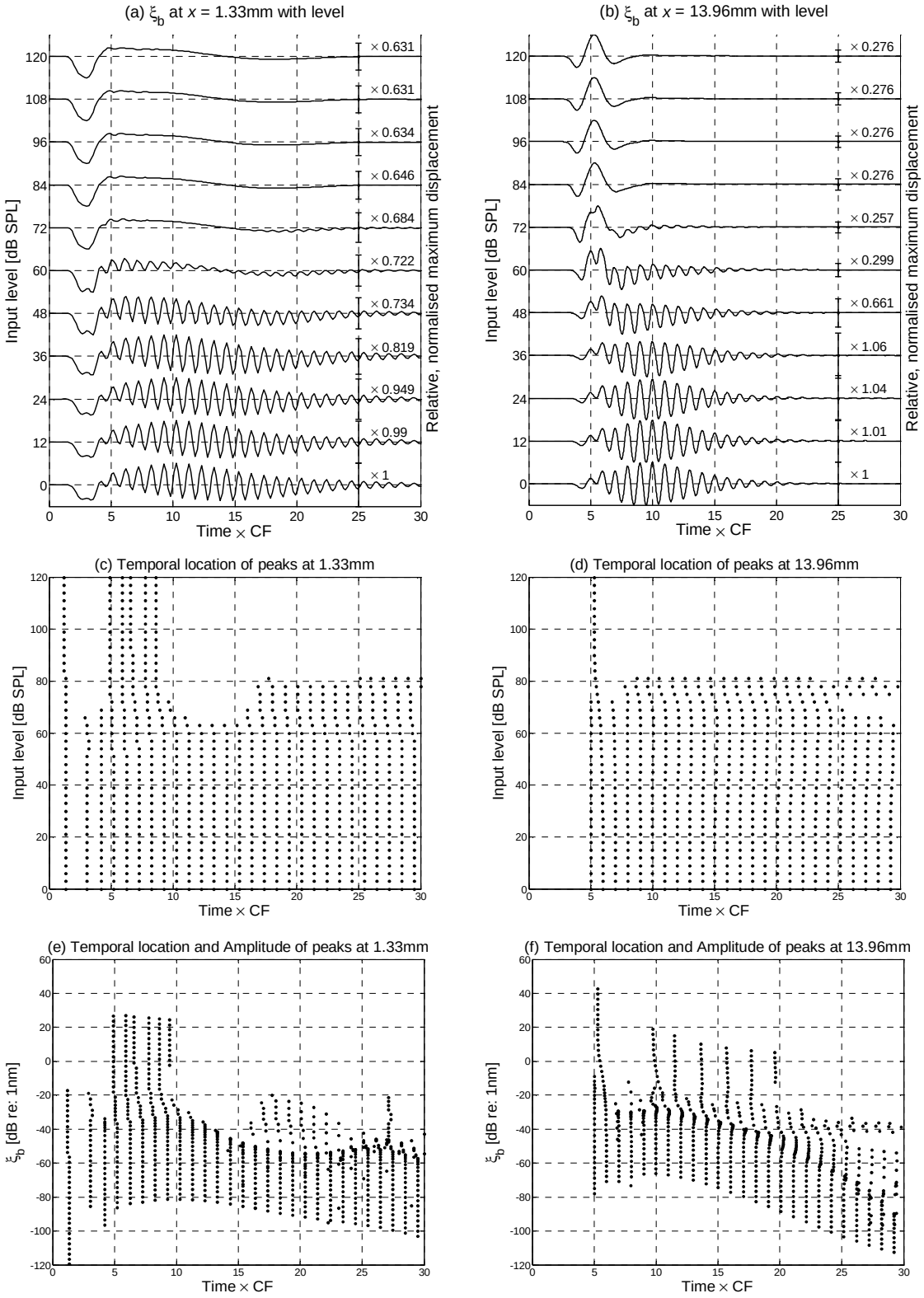


Figure 5.9.a-f: (a - b) show displacement at 1.33 mm and 13.96 mm in a baseline cochlea due to a standard click with variations in stimulus level from 0 dB (bottom) to 120 dB SPL (top). (c - d) plot the locations of the peaks given the same progression from linear to saturating, and (e - f) plot the magnitude of the displacement at the peak with time×CF.

For instance, the oscillations in the fine structure response at moderate stimulus levels persist longer than a linear simulation with a moderate feedback gain. Furthermore, panels (c) and (d) show the peaks of the waveform moving slightly earlier in time with increasing amplitude. This result is different from the analogous linear simulations, shown in Figure 5.3.c-d. Panels (e) and (f) show the same data of (c) and (d), but instead of plotting the results as a function of the driving amplitude, they are plotted as a function of the magnitude of the displacement at the peaks of the waveform. This shows the saturation of the BM responses relative to the temporal changes in the peak location.

Figure 5.10 shows the variation of RMS displacement as a function of stimulus level, calculated over the first 30 cycles of the response at several locations in an active baseline model; the CF at each location was applied to determine the duration of one cycle. The interpretation and speculation regarding these results are left to the discussion at the end of this chapter.

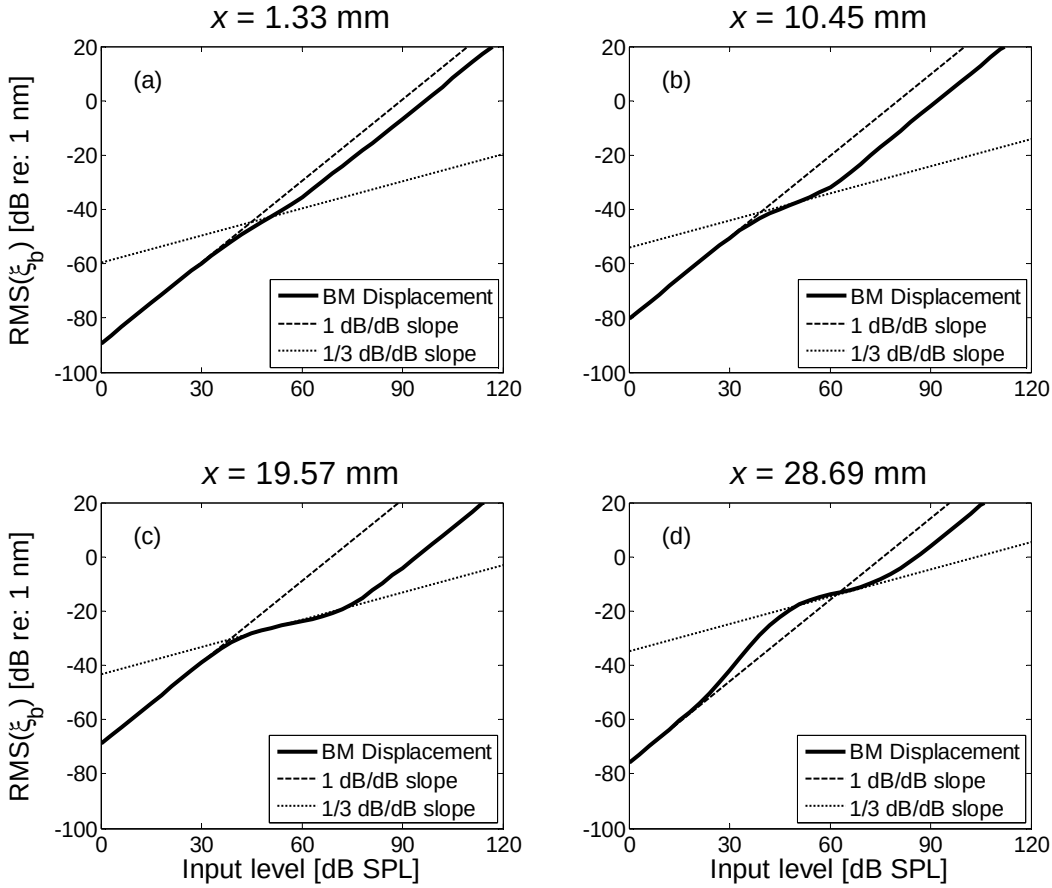


Figure 5.10.a-d: Nonlinear growth of the baseline active model: BM displacement at 4 positions along the BM. RMS values are calculated over the first 30 cycles at each characteristic frequency when $\gamma = 1$.

5.2.2 CEOAEs

As discussed earlier, the latencies of CEOAE frequency components are known to vary inversely with stimulus level. This is predicted in the linear model by the decreasing cochlear group delay as γ is reduced. Similar results are expected in the nonlinear model with effective reductions in γ due to increasing stimulus levels. The results in this subsection are generated from a nonlinear model with the same perturbations in $\gamma(x)$ used in the model of Figure 5.2. A range of click amplitudes was applied, ranging from 0 dB to 120 dB SPL in 3 dB steps. The OAE data was isolated from the stimulus in two ways: 1) directly-determined (DD) by subtracting the basal pressure of the perturbed nonlinear model from the basal pressure of a baseline nonlinear model, as in equation (5.3); and also 2) via the derived nonlinear (DNL) subtraction method,

$$\begin{aligned} DNL\ CEOAE_{st} &= p_{perturbed}\left(amp_i\right)\Big|_{st} - p_{baseline}\left(amp_{i-6dB}\right)\Big|_{st} * \frac{amp_i}{amp_{i-6dB}} \\ DNL\ CEOAE_{ec} &= p_{perturbed}\left(amp_i\right)\Big|_{ec} - p_{baseline}\left(amp_{i-6dB}\right)\Big|_{ec} * \frac{amp_i}{amp_{i-6dB}} \end{aligned} . \quad (5.4)$$

The DD CEOAE is presented in Figure 5.11.a-b as a function of stimulus level. The results are amplitude-normalised by the maximum pressure of each signal in order to more clearly resolve variations in the waveform. The waveform changes little for stimulus levels below approximately 36 dB SPL. From 36 dB to 72 dB SPL, an unusual pulse comes into view at short latencies (1 ms < t < 5 ms), and lower frequency components grow in amplitude at long latencies (t > 25 ms). An abrupt transition in the shape of the waveform takes place between 84 dB and 96 dB SPL where the highest amplitude pressures move from the middle of the time window (~15 ms) to an extremely short delay (t < 1 ms).

To better expose the transition between the moderate- and high-level waveforms, Figure 5.12.a-b shows the normalised CEOAE at amplitudes of 78 dB to 96 dB SPL in 3 dB steps.

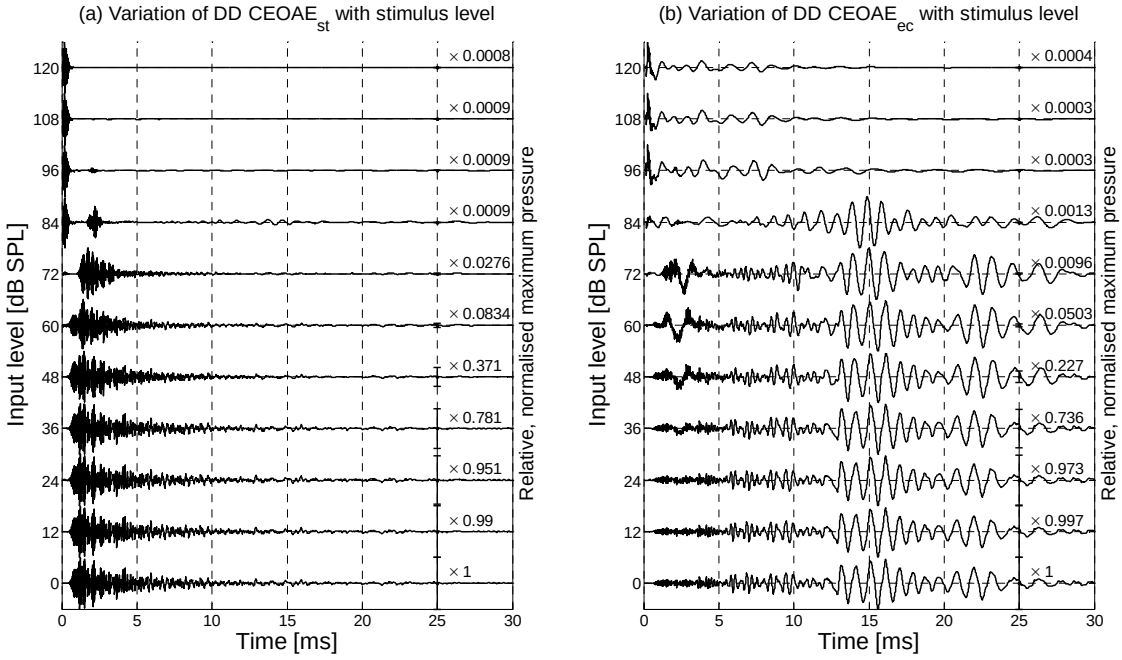


Figure 5.11.a-b: Directly-determined, normalised CEOAE pressure response from the nonlinear model at the stapes (a) and in the ear canal (b) due to a standard click at a wide range of stimulus levels.

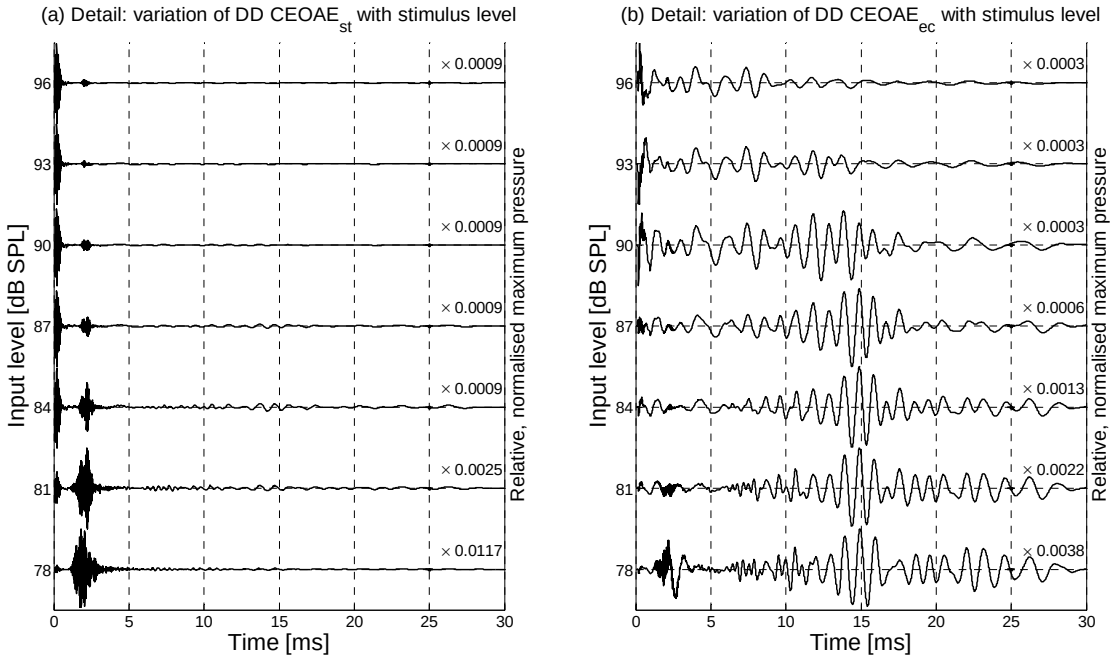


Figure 5.12.a-b: Detailed view of the normalised directly-determined CEOAE pressure response at the stapes (a) and in the ear canal (b) due to a standard click between stimulus levels of 78 to 96 dB SPL.

The CEOAE at the stapes, Figure 5.12.a, shows a gradual changeover in dominance from the oscillations that exist at times > 1 ms (presumably those due to reflections off of the cochlear inhomogeneities) to oscillations that exist at times < 1 ms. The origin of this extremely short-time latency emission (ESLE) is as yet unexplained, though present findings seem to correspond to reports of measured OAEs in humans with extremely short delays at high levels, as discussed at the end of this chapter. In this thesis, the acronym ESLE is adopted to describe any significant signal component that is observed in the 0-1 ms time frame; this can apply to either simulations or clinical measurements of OAEs when the stimulus has been removed.

Variations in the latencies of cochlear reflections are more visible in the ear canal, as shown in Figure 5.12.b. By visual inspection, the relative amplitude of the oscillations at longer latencies decreases and those of shorter latencies increase. The waveform at 93 dB SPL in panel (b) serves as a particularly good 'bridge.' In addition, the higher-frequency oscillations at 5-10 ms at 78 dB SPL are replaced by lower-frequency oscillations over this plotted amplitude range. In order to calculate the CEOAE latencies in each frequency band, the *CWT* method was applied and the first 1 ms of each set of data was discarded to prevent contamination from the ESLE; if this windowing is not performed, the latency at each frequency band jumps discontinuously from low/linear-level latencies to times < 1 ms. Figure 5.13 shows the results of the CWT analysis for the OAE at the stapes (a, c) and in the ear canal (b, d).

While it seems that latencies agree quite well with active cochlear group delay predictions at linear stimulus levels, and do appear to reach the predicted shorter delay at higher (passive) stimulus levels, the transition is quite abrupt in most bands. This is at first rather confusing, as the plots of BM displacement agree fairly well between the baseline linear [Figure 5.3] and nonlinear [Figure 5.9] cases. Further consideration is given in the discussion.

For comparison, the derived nonlinear (DNL) CEOAE of the same data is plotted in Figure 5.14.

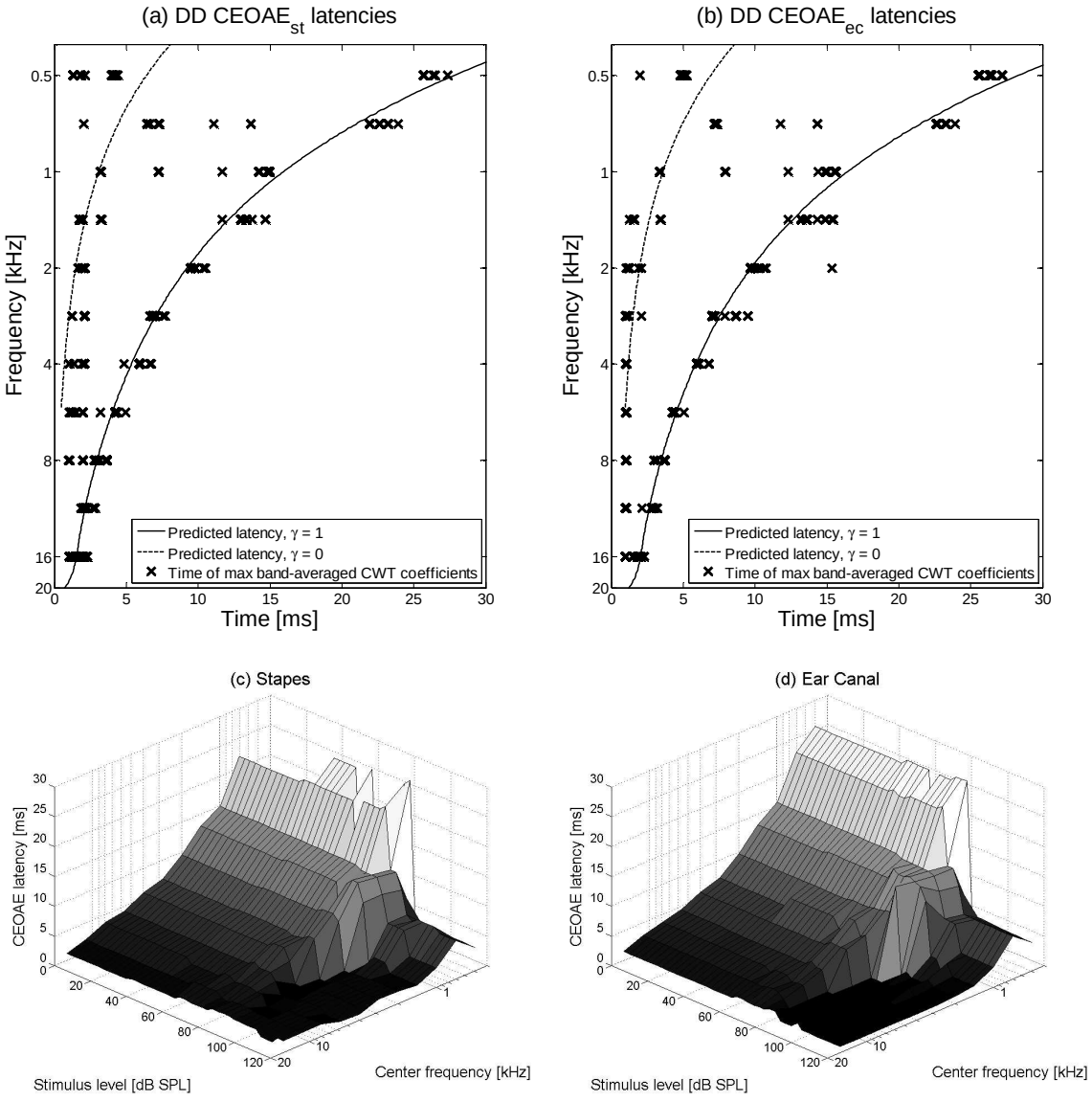


Figure 5.13.a-d: CWT-derived latencies in 11 frequency bands at all stimulus levels. A solid line marks the predicted latency for an active baseline cochlea in (a) and (b), while a dotted line marks the predicted latency for a passive baseline cochlea. (c) and (d) plot the same data, but also against stimulus level.

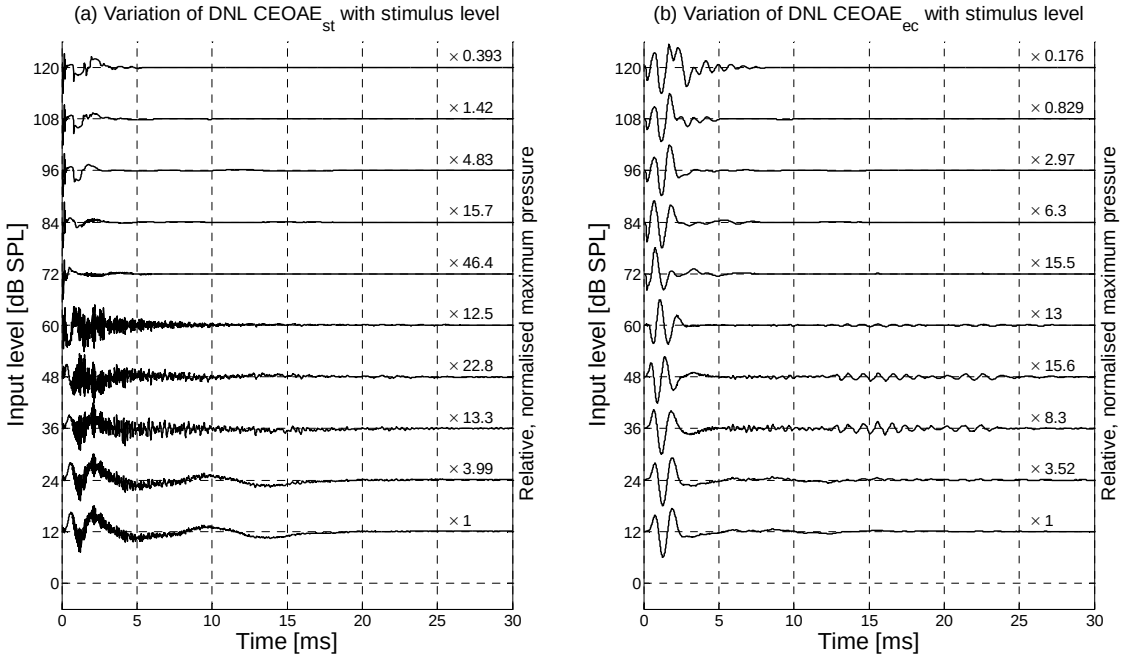


Figure 5.14.a-d: Normalised derived nonlinear CEOAE response at the stapes (a) and in the ear canal (b) due to a standard click with variations in stimulus level.

The growth of the DNL CEOAE agrees with expected behaviour given the expected linear-compressive-linear pattern of BM motion growth as a function of stimulus level. The DNL CEOAE grows rapidly at low-to-moderate levels and then dies away most rapidly in the moderately saturating region of the CA, from 36 dB to 72 dB SPL. However, the finer details are rather unexpected. There appears to be some form of stimulus contamination at each input level, even though the linear components of the waveform have been eliminated. The ESLE begins to dominate the response at higher stimulus levels above 72 dB SPL; this is consistent with observations made of the DD CEOAE waveform as a function of stimulus level. The growth curves of the DD and DNL CEOAEs are presented in Appendix D.

5.2.3 Inhomogeneities in δ(*x*)

It is well established in the literature that place-fixed inhomogeneities can give rise to backward-travelling reflections of TW, at least in models of the cochlea. The source of these inhomogeneities is potentially located in the active process, and has been modelled as a variation in $\gamma(x)$ by Elliott *et al.* (2007). This subsection presents another potential place-fixed inhomogeneity related to the gain; rather than perturbing the gain as a function of

position, the variation is located in the saturating point as a function of position, $\delta(x)$. Physiologically, this might represent ion gates in different sets of OHCs that do not open a uniformly varying quantity.

A consequence of having a perturbation in $\delta(x)$ is that strong reflections should only be detected in the range of moderately saturating stimulus levels. At the lowest levels, one would expect very small amounts of reflection as the gain should be 1 everywhere. At the highest levels, the response of the cochlea is fully saturated (passive) and the saturation point should have little or no impact on the response. Figure 5.15.a-b illustrates the perturbation applied to $\delta(x)$ relative to the baseline distribution. The perturbation, shown in (b), is the same distribution of 'dense' inhomogeneities applied to $\gamma(x)$ in the system of Figure 5.2, but has been scaled such that its new peak-to-peak variation is now 30%.

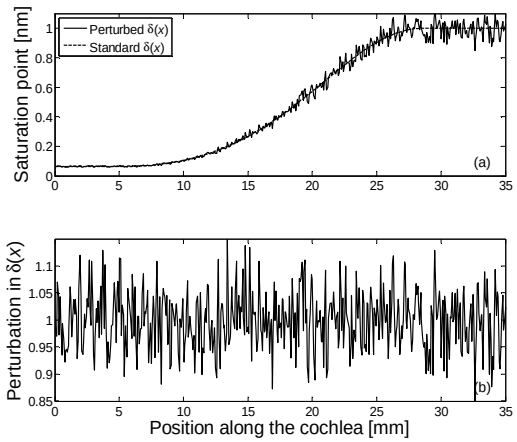


Figure 5.15.a-b: Perturbed saturation point as a function of position. (a) shows the perturbed (solid) and baseline (dashed) saturation points as a function of position. (b) shows the perturbation in isolation.

Figure 5.16 shows the amplitude-normalised CEOAE as a function of stimulus amplitude. It is interesting to note that as the stimulus level is increased, the frequencies that are most strongly expressed vary from high to low.

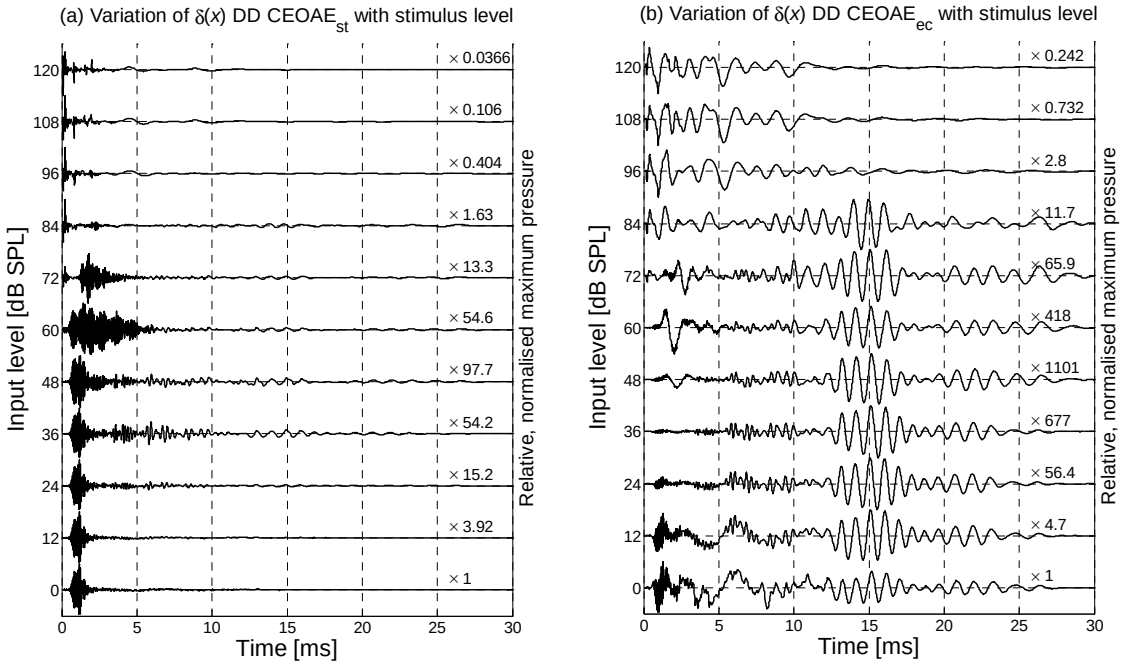


Figure 5.16.a-b: Directly-determined CEOAE in the nonlinear model at the stapes (a) and ear canal (b) due to perturbations in δ(x).

5.3 Discussion

As this thesis is largely concerned with the effects of the variation of feedback gain in the cochlea, it is worthwhile to step back and view the simulations that have been performed with such a perspective. The micromechanical gain in the linear cochlea can be varied, thus setting the base level of feedback at each segment of the BM. The gain can also vary in nonlinear simulations, due to a number of causes. At moderate (but non-saturating) stimulus levels the non-symmetric Boltzmann function decreases the gain for negative relative displacements between the BM and the TM, while it increases the gain for positive relative displacements (see slope of the saturation nonlinearity, Figure 3.15.b). This effect has not been well-studied in this model, and further work is needed in this area. As stimulus levels increase, the active feedback mechanism begins to saturate, thus causing a reduction in effective cochlear amplification. In addition, the saturation point as a function of position, $\delta(x)$, is critical in nonlinear simulations.

As seen in previous chapters, sharp variations in gain can result in impedance mismatches from one portion of the cochlea to the next, thus generating reflections of TWs. This chapter is concerned with the response of the cochlea to clicks, which are by definition a transient stimulus. Except at the very lowest stimulus levels, a linear model

cannot be considered representative of what occurs in the mammalian cochlea. The nonlinear time domain simulation allows $\gamma(x)$ to be varied dynamically which has a direct impact on the motion of the cochlear partition; this is illustrated well by the comparison of the linear and nonlinear BM displacement click responses shown in Figure 5.3.a-b and Figure 5.9.a-b.

One surprising difference between the linear and nonlinear BM responses was the paths of the peaks in time as the amplitude was increased, shown in Figure 5.3c-d and Figure 5.9.c-f. The linear model showed that a reduction in gain produces peaks with increasing latencies, whereas the nonlinear model showed an increase in stimulus level results in peaks with (not always monotonically) decreasing latencies. It is not clear what would produce this effect, though it is likely that amplification of oscillations at moderate levels was occurring at these later peaks in the nonlinear simulation which was not present in the linear case. Such differences between the linear and nonlinear clicks simulations are most visible at moderate levels (e.g. $\gamma = 0.74$, Figure 5.3.a and 60 dB SPL, Figure 5.9.a).

In the nonlinear simulation, the locally active elements saturate near the earliest part of the response, where the amplitude is highest. However, as the response begins to decay away, the lower amplitude waves (of higher frequency) are amplified by the OHCs which are now operating in a more linear range. This effect is understandably absent when examining the corresponding linear plot, Figure 5.3.a, $\gamma = 0.66$ for instance. The gain is constant, and the higher-frequency oscillations decay away much more quickly, whereas they persist in the nonlinear simulation. This is similar to the results of Recio *et al*. (1998), who measured the BM velocity of chinchilla cochleae due to clicks. When the magnitudes of the response at various peaks in time were plotted against stimulus level, different peaks exhibited different saturation characteristics: the earliest and latest peaks showed near-linear growth, whereas the intermediate peaks showed varying amplitude-ranges of compression.

When formulating the saturation point as a function of position, $\delta(x)$, the rationale behind the chosen distribution was that it would allow most of the cochlea to begin to saturate at approximately the same sound pressure level. Perhaps one unintentional consequence of this design was the varying widths of the saturating region of the click-evoked BM displacement growth curves, as seen in Figure 5.10. As a click wave of moderate amplitude travels along the BM, the more basal regions will necessarily be more

saturated than those in the apex as energy in the click is lost. At the most apical regions of the model, it would seem that the range of compressed stimulus levels is greatest; as energy is lost near the base, the OHCs in the middle of the cochlea maintain enhancement of the TW. This, in conjunction with the non-symmetric form of the Boltzmann function, may explain the greater-than-linear growth that is visible in Figure 5.10.d, at 28.69 mm.

One way of testing this hypothesis would be to examine the saturation characteristics of the nonlinear cochlear model to pure tones at several frequencies, and thus characteristic places from base to apex. If the widths of the saturating portion of the response were identical to those of Figure 5.10, then some other explanation would be in order. This is unlikely, however, as the form of the Boltzmann saturation itself is identical at each location; only the saturation point is varied along the BM. Such a response is due to the global nature of cochlear mechanics, and it is difficult to extrapolate this coupled behaviour from response of the micromechanical elements in isolation. Furthermore, these are quantities which cannot be directly measured in human subjects; only OAEs and measurements of BM motion in other mammals are available for study and comparison with model results.

The latencies of the linear CEOAEs generated in this chapter agree well with expectations arising from the cochlear group delays of the model. The average calculated latencies in the set of 100 perturbed cochlear models are within ±5% of predicted latencies. However, the peak-to-peak magnitudes of the simulated CEOAEs are 20-40 dB lower than typically reported results in the literature; this may be due to several factors. The overall levels of BM displacement are lower in this model than is measured in mammals (Robles and Ruggero, 2001); it may simply be that the model should be recalibrated. In addition, the apical saturation point was set at 1 nm in the nonlinear model. This was chosen because it seemed physiologically sensible and resulted in the first signs of CA saturation just below 40 dB SPL, a value which is consistent with measurements in the cochleae of other mammals. If δ was increased, the emitted CEOAE would also be higher in amplitude. However, the most significant factor is likely the magnitude of the perturbation applied to $\gamma(x)$.

The inhomogeneities applied in this chapter were all ‘dense’ and on the order of 0.75%. Increasing this value would undoubtedly result in stronger reflections. The rationale behind keeping this a small amount was to maintain stability. This also serves to

reduce the number of internal reflections such that it would be more straightforward to analyze the results and determine the CEOAE latencies as a function of frequency.

The most surprising results of this chapter concern the nonlinear simulations of CEOAEs. In particular, the rapid change in the form of the CEOAE response over only a 15 dB change in stimulus level, as seen in Figure 5.12, was unexpected. The published variation of CEOAE latencies with amplitude are typically averaged over a large set of results, as in Sisto and Moleti (2007). As such, it is difficult to ascertain whether or not the predicted sharp transition from longer latency to shorter latency shown in the nonlinear CEOAE simulation is normally observed in a single individual. A large set of varied results from many more cochlear models would likely broaden the results, possibly revealing a smoother transition from longer (linear) latencies to shorter (passive) latencies in all frequency regions. It is also possible that perturbations in the overall saturation level and perhaps even different configurations of $\delta(x)$ would result in smoother variation of latencies when averaged over a large set of results. Another unexpected result of the nonlinear CEOAE simulations was the prominence of an extremely short latency emission (ESLE) at high stimulus levels which was presented at times < 1 ms (see Figure 5.11).

Other researchers have observed phenomena similar to the ESLE in these simulations while measuring various OAEs at moderate- to high- levels; a variety of explanations for their source have been offered. Nonlinearities in the transducer producing the stimulus (Konrad-Martin and Keefe, 2005), distortion in middle-ear transmission possibly due to the acoustic reflex (Konrad-Martin and Keefe, 2003; Guinan *et al*., 2003), and 'fast' compression waves in the cochlear fluid (Ren, 2002) have all been proposed. As none of these mechanisms are present in the present model, they cannot be the cause of the ESLE in these simulations. The most likely explanation is the existence of a reflection or distortion site located at or near the base of the cochlea (Brass and Kemp, 1993; Talmadge *et al*., 2000; Konrad-Martin and Keefe, 2005). However, it is also possible that the source of the ESLE in these simulations does not correspond to the source(s) of the ESLEs in clinical measurements. For instance, there may be a number of sources in measurements that combine to give rise to an ESLE. The simulated growth curves of the CEOAEs in the nonlinear model, calculated at different time frames, may shed some light on this discussion; these are included in Appendix E.

Finally, it should be noted that the simulations concerning perturbations in $\delta(x)$ should not be taken to suggest the author is proposing that this is a dominant property of the cochlea. The fact that variations in ear canal pressure within a local band of frequencies are most visible at low stimulus levels suggests that the reflection mechanism is linear (e.g. Zwicker, 1990). Nevertheless, it is an interesting experiment to run as a comparison to the CEOAEs generated given inhomogeneities in $\gamma(x)$. For instance, certain low, medium, and high frequencies are expressed more strongly at high, moderate and low stimulus levels, respectively, in Figure 5.16.a-b. This suggests that the cochlear model saturates first at the base and then toward the apex because the perturbations in this simulation are present in $\delta(x)$.

The ability to apply various perturbations to the cochlear model and to compare the resultant BM motion to 'known' or baseline responses is of great use. The power of the nonlinear time domain simulation of the cochlea is that it allows the modeller to probe into many areas of as-yet unexplored cochlear function. It is clear, however, that of the nonlinearity as a function of position in the cochlea is yet another aspect of the model that requires careful consideration and tuning.

Chapter 6

Conclusions and Suggestions for Further Work

This chapter summarizes the strengths and weaknesses of the state space model for modelling measured and inferred features of the human cochlea. As a result of this work, numerous potential research directions have also become apparent. These are discussed in subsection 6.3.

6.1 Strengths of the state space model

The original Neely and Kim (1986) frequency domain model was among the first mathematical representations of a mammalian cochlea that included an active element and a second degree of freedom in the micromechanics. These attributes allow for amplification of tonally excited TWs at different positions along the cochlea. In addition, the Neely and Kim (1986) formulation is based on a physical interpretation of the CP's micromechanics that provides plausible results. There are many other such models in the literature. Another well-referenced representation of the cochlea is Zweig's (1991) delayed-stiffness model, for example. However, the inclusion of a delay in the system's dynamics leads to some analytical complications. A lumped parameter system, such as the Neely and Kim (1986) model, is described by an ordinary differential equation which generates a finite number of poles. In contrast, a system that includes a delay can generally be described by a partial differential equation and generates an infinite number of poles (Franklin *et al*., 1991). This is undesirable in the context of this thesis because the interpretation of results becomes less straightforward.

One downside of the original Neely and Kim (1986) model was that the nominal feedback gain of the original formulation lead to a much higher enhancement of the BM response than has been measured *in vivo* (Hubbard and Mountain, 1996; Robles and Ruggero, 2001). This and a number of other characteristics of the system were modified in order to represent a human cochlea. By reformulating the model from a frequency domain representation to a state space representation, it also became possible to perform linear and nonlinear time domain simulations. The responses of the model that can be compared with experimental data and observations are discussed below.

6.1.1 Cochlear responses

The revised parameters presented in this thesis are able to account for the following features observed in the mammalian cochlea:

- good fit to the human frequency-to-place map at all but the most apical locations;
- BM enhancement curve that is approximately 45 dB at the base and 20 dB at the apex;
- basalward shift in the location of the peak of the TW given reductions in feedback gain; equivalently stated, the maximum response at a single position along the coupled cochlea shifts to lower frequencies given reductions in feedback gain;
- variation of λ_{peak} with position along the cochlea that is consistent with inverse calculations made from clinically collected OAE data (Shera and Guinan, 2003);
- both harmonic and intermodulation distortion at various positions along the nonlinear time domain model of the cochlea, similar to what is observed in the biological cochlea;
- compressive growth rates of BM motion at moderately saturating stimulus levels;
- a basal boundary condition that is similar to the impedance looking out of the cochlea and into the middle ear.

Furthermore, the active portion of the cochlear model is approximately 2 mm wide for a given frequency, basal to the characteristic place. This is a value that has support from experimental inferences (Allen and Fahey, 1992).

One of the primary strengths of the state space model is its ability to quickly and unambiguously determine the stability of the linear model. This is important to validate frequency responses of the linear system and also has implications for the study of SOAEs.

6.1.2 SOAEs

The linear state space model exhibits features that are consistent with the multiple-reflection theory of Zweig and Shera (1995). In this theory, the frequency-to-place map and the variation of the wavelength of the TW at its peak dictate the spacings between linear instabilities given a 'dense' set of reflection sites along the BM. The gain of the CA, $\gamma(x)$, was chosen as the parameter to be perturbed in this investigation. The spacings between adjacent unstable frequencies were found to match theoretical predictions only when the inhomogeneities were 'dense.' However, fewer and less strongly undamped instabilities were generated when 'sparse' perturbations in $\gamma(x)$ the same magnitude were applied.

The ability to compare the linear stability of a cochlear model with its nonlinear limit cycle behaviour is one of the strengths of this model with regard to SOAEs. The investigations presented in Chapter 4 evaluated the pressure at the base of the cochlea at varying time frames both soon and long after the initial stimulus had begun to decay away. In the nonlinear simulations of Section 4.3.1, there was only a single linear instability. The steady state frequency of both the dominant BM oscillation and the pressure at the base differed from that of the unstable pole by less than a tenth of a percent. Components of the BM motion at the frequencies of the near-unstable poles were initially observed, but fell away into the noise floor within several hundred milliseconds. The $2f_0$ and $3f_0$ harmonic components of the near- and fully-unstable frequencies were present so long as the response at the primary tones remained strong; however, the magnitudes of the distortion frequencies observed at the stapes were typically 40 and 80 dB below the primary signal, respectively.

The magnitude and phase of the BM velocity as a function of position along the cochlear model of Section 4.3.1 were calculated at the limit cycle frequency. At the fundamental frequency, there was a single position where the TW was 'standing,' as indicated by $\partial\phi/\partial x = 0$. This was located slightly basal of the peak in the response, in the region of negative damping. This phase characteristic indicated that the amplitudes of the forward- and backward- TWs were balanced in this area. The magnitude of the response appeared to show that both the forward- and the backward-TWs were amplified, similar to simulation results produced by Talmadge *et al*. (1998). In conclusion, the forward TW is

amplified basal to the peak region, and its reflection is also amplified when heading back toward the stapes. This resulted in the division of the cochlea into two portions: one basal, where the TW was dominated by backward propagation, and one apical, where the TW was dominated by forward propagation.

The global reflection theory of Shera and Zweig (1993) and Zweig and Shera (1995) was developed to explain the commonly observed spacings between SOAE frequencies. This linear theory is well-supported by the stability predictions of the linear state space model. However, the nonlinear model demonstrated that limit cycles can interact to suppress one another in Section 4.3.2. Only three of the five linearly unstable frequencies were strongly expressed as limit cycles at the end of a 3000 ms simulation. Numerous oscillations at other frequencies also appeared as a result of both harmonic and intermodulation distortion. Two of the linearly unstable limit cycles approached a steady state amplitude within the first 200 ms of simulation. However, most of the magnitudes of these frequency components varied as a function of time.

The near-unstable frequency components decayed at the steepest rates. This finding is in good agreement with the experimental results of Sisto *et al.* (2001), who show the presence of both 'long-lasting' OAEs and exponentially decaying components of sharp tuning in humans. However, whereas they conclude that the presence of both of these responses suggests that 'OAE dynamical properties are mainly determined… by the local cochlear parameters, rather than by the overall cochlear transmission,' the simulations presented in Chapter 4 show that the global reflection of TWs can explain both phenomena.

The distances between the frequencies of the detected limit cycle components were very regular; this was primarily believed to be due to the nature of the intermodulation distortion generated by the 'primary' tones at the linearly unstable frequencies. There have also been reports of measured SOAEs that are the result of intermodulation distortion (Burns *et al.*, 1984; Whitehead, 1988). Whether the regular spacing between simulated limit cycle oscillations is detectable in the ear canal after reverse-transmission through the middle ear, and if it is still present locally when many more linear instabilities are present, is very much an open question. Such behaviour is also qualitatively similar to the nonlinear effect of mode-locking in a self-excited system, such as a wind or string

instrument (Fletcher and Rossing, 1998). We may one day learn that the appreciation of tonality in music is very much connected with the nonlinearities of the cochlea.

6.1.3 CEOAEs

In Chapter 5, CEOAEs were generated by sending clicks into densely perturbed cochlear models. Commonly measured features of CEOAEs were compared against simulation results. Perhaps the most satisfying finding of the investigation into CEOAEs concerned the frequency-dependent latencies of the reflections. Linear model results matched the predictions of the group delay of the model and were very similar to the most commonly measured delays in clinical measurements. This outcome is perhaps attributed to the direct relationship between the wavelength of the TW and its wavespeed, as given in equations (4.7) and (4.8). Thus, working backwards from SOAE spacings as a function of frequency to develop a distribution of $\lambda_{peak}(x)$ also set the correct group delay in the cochlear model.

A more conventional finding was that the linear model agreed with other studies (e.g. Sisto and Moleti, 2002; Sisto *et al.*, 2007) which show that the frequency-dependent CEOAE latency is approximately twice the forward travel time of the TW. However, linear estimates of the delay were somewhat longer than what is commonly reported in the literature (Tognola *et al*., 1999).

In most clinical CEOAE experiments, it is necessary to set the level of the stimulus clicks at a minimum of 60 dB SPL in order to achieve an acceptable signal-to-noise ratio (Hall, 2000). However, this amplitude is within the saturating region of the CA. As shown by Figure 2.21, the cochlear group delay decreases as the feedback gain is reduced; this is analogous to the effect of increasing the stimulus level of the signal. Thus, it is expected that CEOAE delays calculated in the linear model will be longer than published experimental data.

The nonlinear CEOAE simulation showed a decrease in latency within each frequency band was given increasing stimulus levels. However, there were some frequency bands that showed an abrupt decrease in the latency with level. It is difficult to gauge the generality of this single model. The averaged results of a larger number of cochlear models may show a smoother transition of level-dependent latencies more similar to that reported in human data, as collected by Sisto and Moleti (2007) for example.

The last finding of interest with regard to the nonlinear CEOAE simulations was the discovery of an extremely short latency emission (ESLE), qualitatively comparable to signals detected within several ms of a moderate-to-high level stimulus as reported in the literature (e.g. Konrad-Martin and Keefe, 2005). The most likely explanation for the ESLE in the simulated- and perhaps the measured- results is the existence of a reflection or distortion site located at or near the base of the cochlea (Brass and Kemp, 1993; Talmadge *et al*., 2000; Konrad-Martin and Keefe, 2005). However, the growth curves of CEOAE amplitudes, found in Appendix D, show that the ESLE grows linearly given perturbations in $\gamma(x)$ at high amplitudes. This would suggest that a distortion explanation is unlikely; such phenomena are typically compressive in nature. Although the ESLE waveform grows linearly in the model with inhomogeneities in $\gamma(x)$, it does not in the model with perturbations in $\delta(x)$. This suggests that slightly different mechanisms may be operating here. This is still somewhat perplexing; one would expect both forms of gain-based perturbation to become suppressed at the highest levels. Other nonlinear simulations not presented in this thesis suggest that this effect is not due to the particular distribution of $\gamma(x)$ applied in Chapter 5.

6.2 Weaknesses of the state space model

Though it has been shown that the state space model is capable of exhibiting many of the salient features of the human cochlea, there are a number of weaknesses that one should be aware of. These are categorised and discussed in order of decreasing prominence.

6.2.1 Cochlear responses

Perhaps the most significant shortcoming of the Neely and Kim formulation is the manner in which its CA is modelled. Early attempts to represent the micromechanics included an active element that shifted the peak in the element's admittance in frequency by varying its stiffness. This was implemented to describe the half-octave shift observed along the tonotopy of the BM given quiet vs. loud sounds. However, what early modellers did not appreciate is that the location of the amplified TW peak does not shift apically from the passive TW peak as a result of variations in the micromechanical tuning. Instead, the amplification takes place at the local resonant frequency of the CP.

The shift in the location of the TW peak is a result of the amplification of the forward TW, which increases in amplitude as it propagates through the negative damping region. The TW then peaks approximately 1 to 2 mm apical of the negative damping region. As Shera (2001b) shows, the intensity-invariance of the fine time structure in BM click responses is closely related to the tuning of the micromechanics as the gain is linearly varied. However, results presented in Chapter 5 showed that the temporal location of the peaks in the fine structure of nonlinear BM click responses can vary non-monotonically, unlike linear predictions. This appeared to be primarily due to the fast recovery of the CA on a cycle-by-cycle basis, depending on the amplitude of the stimulus. Unfortunately, this behaviour is not observed *in vivo* (Recio and Rhode, 2000; Robles and Ruggero, 2001), and thus is likely just to be a feature of the current model.

The active impedance of the Neely and Kim (1986) model is also greatly simplified. None of the more complex mechanics of the OHC are taken into account, such as the stiffness of the ion gates or any local fluid viscosity. The magnitude of the model's active force will also grow without limit as the driving frequency increases, whereas it has been shown that this is not the case in isolated OHCs. (Kros, 1996)

In summary, the current form of the CA in the state space formulation presents the most serious drawback of the model with regard to reproducing measured phenomenon in the mammalian cochlea. However, this fact should not be taken to suggest that investigations performed with this model are completely invalid or without worth. As shown above, the state space model is capable of exhibiting a wide variety of features of the human cochlea. A model is by definition a simplification of a more complex system, and is only capable of describing a subset of its properties. Thus, one could observe that the state space model best represents the cochlea at its baseline active state and at lightly saturating levels in the nonlinear regime. Simulations outside of this range of operation can still provide insights into the behaviour of the biological cochlea, but care and consideration must be taken when drawing conclusions. In addition, there are yet further enhancements that could potentially improve the accuracy of the model's response.

For instance, it is possible to reduce higher-dimensional representations of the fluid dynamics in the cochlea into a form usable by a 1-D box model (Mammano and Nobili, 1993). This is desirable because the long-wave assumption is violated near the peak of the TW, where the wavelength of the TW is on the same order as the dimensions of the

cochlear cross-section. This introduces the possibility of a resonance in the transverse direction, or a 'short-wave' mode. In the literature, the matrix that describes the coupling of motion from one position along the cochlea to another is called the Green's function; the equivalent quantity in the state space model is the $\mathbf{F}^{-1}$ matrix. The primary difference between the 1-D and the 3-D Green's functions is an added sharpening in the local coupling at nearby locations. This can also serve to modify the phase near and beyond the peak of the TW in such a way that is more similar to measured responses (Kolston, 2000).

However, as discussed by Shera *et al*. (2005), adding higher dimensional approximations do not appreciably affect the fundamental relationships internal to the model, such as the connection between TW group delay and its wavelength. One important effect of adding higher dimensionality is that the reflections from the inhomogeneities in the peak region are also enhanced; this would make the system even more prone to instability arising from inhomogeneities along the CP, without modifying the existing active element in the micromechanics.

A common criticism of early cochlear models, particularly that of Neely and Kim (1986), is that the active pressure generated by the OHCs have nothing to react off of (e.g. Hubbard and Mountain, 1996). This represents a serious shortcoming in terms of relating the model's response back to the local motion of the CP. A number of authors have suggested that OHCs may react off of adjacent segments of the BM, thus producing a so-called mechanical 'feed-forward' coupling (Kolston *et al*., 1989; Steele *et al*., 1993; Geisler and Sang, 1995; Fukazawa, 2002).

One final comment on the cochlear response of the nonlinear model concerns the growth rates of the BM motion due to both transient and steady-state stimuli. While compressive growth rates have been calculated at moderately saturating levels, a number of results seem to show greater-than-linear growth at near-saturating levels. The $2f_0$ growth curve shown in Figure 3.20 is in fact similar to the equivalent result in Figure 3A of (Cooper, 1998); however, the $2f_0$ magnitude exceeds that of the fundamental in the simulation presented in Chapter 3, whereas it does not in Cooper's (1998) experimental measurement. In addition, the growth curves of the click-evoked BM response in Figure 5.10.d, and also those of the RMS CEOAE pressure in Appendix D, are again greater than linear. This is in disagreement with numerous OAE and BM measurements (e.g. Robles

and Ruggero, 2001). The source of this mysterious result may be the form of the Boltzmann nonlinearity in the model.

The values of the Boltzmann function chosen to saturate the input to the feedback force give rise to a non-symmetrical input-output curve, as shown in Figure 3.15. Figure 3.15.b illustrates that the slope of the nonlinearity is unity at small input values. This ensures a perfectly linear response at low stimulus levels, as shown by the BM responses of Figure 3.18 and Figure 3.20. At moderate input levels, the slope of the nonlinearity becomes non-symmetrical and actually exceeds unity for positive inputs. This effectively turns up the gain (for positive inputs) and may explain the greater-than-unity growth curves mentioned above.

6.2.2 OAEs

Many aspects of the OAEs simulated by the model in this thesis match experimentally measured results (i.e. spacings between linear instabilities and frequency-dependent CEOAE latencies). However, the magnitudes of the simulated OAEs are typically 20-40 dB lower than equivalent published results. This is very straightforward to remedy in the simulation of SOAEs, as generated by limit cycle oscillations in the nonlinear cochlea. Increasing the saturation point of the nonlinearity in the CA by a factor of 10, for example, would cause a corresponding +20 dB shift in the magnitudes of the limit cycle oscillations. This modification has support from the nonlinear simulations of BM growth curves due to a 3 kHz tone, as shown in Figure 3.20.a. This figure shows that the fundamental response transitions from low-level (linear) to moderate-level (compressive) growth at approximately 5 dB SPL; in experimental results of BM motion, this first 'corner' in the growth curve is typically measured at ~30-40 dB SPL (Cooper, 1998; Robles and Ruggero, 2001). Thus, by increasing the saturation point by a factor of 10 (or more), the 'corner' in the growth curve would shift up and come more in line with experimental measurements.

The magnitudes of the linear results are somewhat more difficult to correct. One contributing factor to the offset may be the assumed input impedance of the cochlea when calculating the stapes acceleration given a volume displacement in the ear canal. This was set at a flat value of $1.1*10^{10}$ Acoustic Ohms early in the work, before the model was revised. Current calculations (see Figure 3.11) show that this is a slight overestimate of the actual cochlear input impedance. However, this correction would be a minor improvement

at best. Another potential solution is the inclusion of a short-wave resonance, which would necessarily magnify the response at the peak and thus the energy reflected at the peak. However, this may also be a more basic shortcoming of the model parameters that could be addressed in the future.

6.3 Suggestions for future work

Many potential areas of future work have become apparent through the course of this doctoral investigation. This subsection is divided into three areas: 1) further tuning and study of the model; 2) further simulations of cochlear phenomena; and 3) wider topics of research.

6.3.1 Further study and tuning of the model

The most straightforward and pressing aspects of the model that require study pertain to the form of the saturation nonlinearity. Early nonlinear simulations applied a simpler hyperbolic-tangent function to saturate the feedback force (Elliott *et al*., 2007). A drawback of the hyperbolic tangent function is that it is symmetrical, and thus is incapable of generating the even-order harmonics that are commonly measured in the BM motion (Cooper, 1998). The Boltzmann function was applied to the state space model here, because of its similarity to measured input-output characteristics of OHCs in isolation (Cody and Russel, 1987; Kros *et al*., 1992); this followed the precedence of other work which has also relied upon the Boltzmann equation to describe the saturation of the OHC feedback force with increasing stimulus level (e.g. Nobili and Mammano, 1996).

As noted above, the Boltzmann function is a description of the mechano-electrical transduction characteristics of an OHC in isolation. However, the force generated by the OHCs in the organ of Corti *in vivo* is not well characterised. This discussion is important in the context of the Neely and Kim (1986) framework, as the details of the chemical and mechanical OHC dynamics are hidden by representing the active element as an impedance. This requires some tuning the Boltzmann function in the nonlinear state space model to account for observed cochlear responses. A detailed study of the effect of variations in the saturation function upon the simulated nonlinear cochlear response has yet to be performed. A logical starting point would be to analyse the response of an isolated nonlinear micromechanical element. The variation of the Boltzmann function parameters

that control its asymmetry and slope would necessarily change the model's response. The relative magnitudes of the fundamental and harmonic components may also provide important insight into how distortion propagates in the coupled cochlea. Finally, tuning the saturation point as a function of position, $\delta(x)$, may better match measured results.

Another aspect of the model that has not been studied is the power gain produced by the CA. The enhancement of the TW is often calculated because it is directly observed by experimentalists. However, it may turn out that the power output of the individual hair cells is not physiologically plausible, as in the original Neely and Kim (1986) formulation (Hubbard and Mountain, 1996).

A more mundane but equally important area of tuning the model concerns the logistics of the time domain simulations. The primary limiting factors of performing extended simulations are computational time and memory (both random-access memory and read-only memory). Late in this Ph.D., it was discovered that fine-tuning the error tolerances for each individual state can reduce simulation times by fourfold without sacrificing the accuracy of the results. Additional efforts in this area may yield further reductions in computational load, perhaps along the lines of Diependaal's (1987) work which applied a time-varying spatial discretisation map. Initial attempts to improve efficiency by pre-compiling the MATLAB function in C language have proved largely unsuccessful; this is believed to be due to the computationally intense nature of the ODE solver.

At the time of writing, the preferred method of running extended simulations is to break up the simulation into smaller blocks of time, typically 100 ms. This is accomplished by setting the initial values of the states of the new simulation to the values of the states at the end of the previous simulation. This is necessary as the 32-bit architectures and Microsoft Windows operating systems employed by most computers effectively limit the (virtual and real) RAM usage of MATLAB variables to approximately 1.2 GB. However, this has been overcome with the introduction of 64-bit architectures (and a version of MATLAB capable of using a 64-bit system) where the primary limitation is the physical storage space available. (Mathworks, 2008)

6.3.2 Calculations and simulations

Most of the suggestions presented here concern the nonlinear response of the state space model, as linear systems have been well-studied and can only represent a very limited range of cochlear phenomena. The most promising and exciting line of inquiry following this work concerns the nonlinear interaction of linear instabilities, the initial results of which were presented in Chapter 4.

Combining linear theories of cochlear function, such as the coherent reflection theory of Zweig and Shera (1995), with nonlinear simulation will likely reveal something of the deeper character of the cochlea. The results of Chapter 4 showed that densely perturbed cochlear models produce instabilities at frequencies which share regions of negative damping. However, in the presence of other nearby instabilities, the magnitude of a linear instability's undamping ratio does not necessarily seem to map well to the magnitude of the steady state nonlinear limit cycle oscillation. An interesting simulation to run would be to count the number of limit cycles detectable in the ear canal above the experimental noise floor, and to compare this value against the number of linear instabilities. This area of work requires the study of nonlinear suppression, where one tone can affect the response of another.

Early in this Ph.D., the suppression characteristics of Van der Pol Oscillators were briefly simulated. Some of these initial results are included in Appendix E. It is hoped that comparisons between the suppression of Van der Pol oscillators and the suppression of nonlinear limit cycle oscillations will offer some clues as to the nature of the suppression in the cochlea. It has also been demonstrated by numerous experimentalists (e.g. Zwicker and Schloth, 1984) that externally applied excitations can frequency-lock, phase-synchronize, suppress, or otherwise affect a SOAE. These phenomena would be interesting to study in the nonlinear state space model, and could potentially be compared against the results of Sisto *et al*. (2001) who studied the transient dynamics of click-synchronised SOAEs. This would also tie in with the simulation of CEOAEs generated in the presence of instabilities, which has yet to be simulated in the state space model.

The simulations of CEOAEs revealed a number of unexplained results, such as the presence of an extremely short latency emission at higher stimulus intensities. Similar results have been noted in the literature (Brass and Kemp, 1993; Talmadge *et al*., 2000; Ren, 2002; Guinan *et al*., 2003; Konrad-Martin and Keefe, 2003; Konrad-Martin and

Keefe, 2005), but no satisfactory generation mechanisms have been agreed upon. A variety of different stimuli and variously-perturbed models may help determine if the ESLE is a nonlinearly distorted reflection of the input, or due to some other mechanism. Further nonlinear simulations of CEOAEs are also necessary to determine if the sharp transition from long to short latencies with increasing amplitude observed in the current model persists when a large set of results is averaged.

The dominant mechanism for the generation of SOAEs and CEOAEs in this model appears to be reflection; this is in accordance with the taxonomy of OAEs described by Shera and Guinan (1999). While the nonlinear state space model clearly shows evidence of distortion, a detailed study of DPOAEs and other distortion-related phenomena has not yet been undertaken. Such simulations may reveal some finer details of the wave-fixed mode of DPOAE generation and propagation. In anticipation of this work, it would be instructive to simulate SFOAEs in order to ground the research.

Lastly, a more rigorous approach to quantifying the reflection in the cochlear model due to inhomogeneities would help clarify results. This could be accomplished by decomposing the pressure and motion at each position into to forward- and backward- TW components, perhaps using the WKB-method as applied by Zweig (1991) or Neely and Allen (2008). There are, however, difficulties concerning the underlying assumptions of this method in a nonlinear system such as the cochlea.

6.3.3 Wider topics of research

Within the field of cochlear modelling, the Neely and Kim (1986) framework represents a compromise between a completely phenomenological representation and an ultra-detailed 3D finite element model of the cochlea. The model is not so complex that more detailed investigations become prohibitively expensive in terms of computation time, and yet it based in the physics of the cochlea. For this reason it is well-suited to study the nonlinear characteristics and interactions of TWs along the CP, given the computational limits at this time.

The propagation of a second TW mode along the TM is another area of study that is becoming more widely studied in the field of cochlear mechanics (Hubbard and Mountain, 1996). Whether longitudinal coupling through the TM in a model as reduced as the Neely and Kim (1986) formulation would produce any worthwhile results is uncertain. However,

the system studied in this investigation and its various lumped element parameters do have direct correlates to the physical structure of the biology. An important area of future work consists of further collaboration with experimentalists to determine physiologically plausible parameters for cochlear models. It is also possible that simulations with complex, three-dimensional finite element models based on physiological measurements (e.g. Meaud and Grosh, 2008) may suggest more appropriate parameters for simplified models. This would be particularly useful if the goal is to simulate a physical abnormality in the cochlea, sensorineural hearing loss for instance.

The author strongly believes that the mechanical modelling of cochlear pathology is a crucially important research focus. The mammalian cochlea is a fascinating system to study, full of unexpected nonlinear phenomena and complexities. Although there is still a great deal of basic research to be done on this sensory organ, it is important to be mindful of the wider implications of its study. For instance, more than 8% of the population of many developed countries suffer from significant sensorineural hearing loss. In addition, approximately 90% of all hearing loss in adults is due to cochlear malfunction (Jesteadt, 1997). Thus, it is hoped that the work presented in this doctoral thesis may begin to provide some insight into the inner workings of the cochlea for any researchers interested in studying the mechanisms of hearing loss in the future.

Appendix A

Middle Ear and Ear Canal Model

The primary function of the middle ear, as shown in Figure A.1, is to match the relatively low impedance of the air in the ear canal to the relatively high impedance of the cochlear fluid, thus ensuring efficient transfer of acoustical energy. However, the middle ear and ear canal each add their own signature to the forward- and reverse-transmission of sounds; these characteristics also impact the middle ear boundary impedance at the base of the cochlea. In this appendix, the forward- and reverse-transmission characteristics of a middle ear and ear canal model are illustrated in the context of frequency- and time-domain simulations. In addition, the impedance of the middle ear and ear canal as measured from the cochlea is modelled in a manner such that it can be easily incorporated into the state space model.

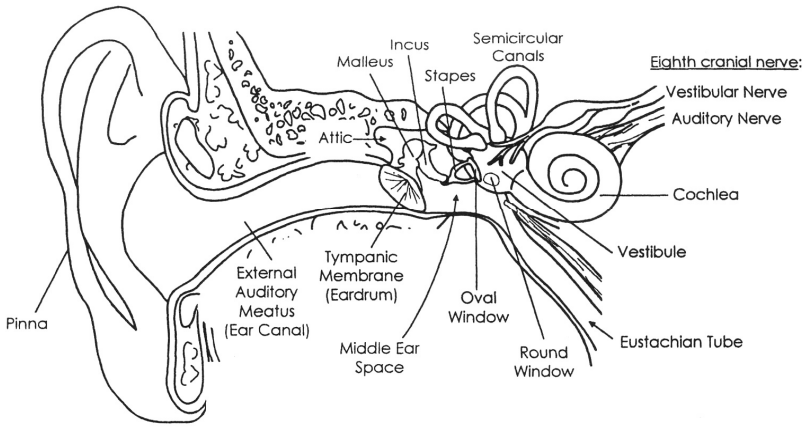


Figure A.1 The human auditory system. Note this figure is not to scale. Reproduced from 'Hearing: an introduction to psychological and physiological acoustics,' by S.A. Gelfand, Copyright (1998), with permission from Marcel Dekker.

A.1 Two-Port Networks

Frequency domain models of the middle ear and ear canal are often cast in the form of two-port networks that describe the mechanics of the system (e.g. Kringlebotn, 1988). The two transmission matrices for the middle and inner ear can be cascaded to produce a single transmission matrix; this overall matrix then relates the input acoustic pressure and acoustic volume velocity at the stapes, $P_{st}(\omega)$ and $Q_{st}(\omega)$, to give an output acoustic pressure and acoustic volume velocity at the ear canal, $P_{ec}(\omega)$ and $Q_{ec}(\omega)$. This is illustrated in Figure A.2, and formalised in Equations (A.1-3). It should be noted that the pressures and volume velocities at the stapes and ear canal and all transfer matrices are functions of frequency unless otherwise noted; the explicit notation of this is suppressed for convenience here.

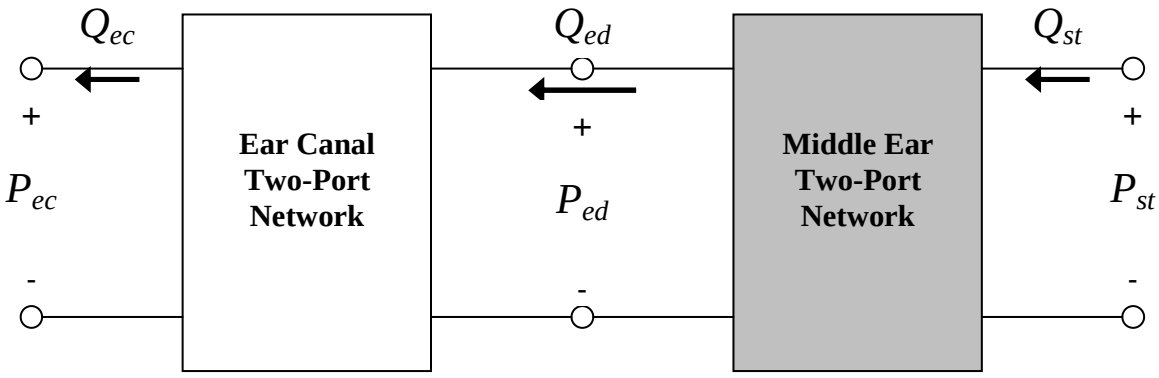


Figure A.2: Two-port network representations of the ear canal and middle ear, where *P* and *Q* are pressure and volume velocity, respectively, at the ear canal (*ec*), eardrum (*ed*) and stapes (*st*).

The behaviour of the middle ear is characterised when the acoustic pressure and volume velocity at the eardrum is expressed in terms of the pressure and volume velocity at the stapes:

$$\begin{bmatrix} P_{ed} \\ Q_{ed} \end{bmatrix} = \begin{bmatrix} T_{edst11} & T_{edst12} \\ T_{edst21} & T_{edst22} \end{bmatrix} \begin{bmatrix} P_{st} \\ Q_{st} \end{bmatrix}. \tag{A.1}$$

Note that the elements of the transmission matrix must be defined such that the volume velocities are all travelling in the same direction; this allows multiple two-port models to be readily cascaded as in a transmission line.

The output at the end of the ear canal is expressed given the input at the eardrum:

$$\begin{bmatrix} P_{ec} \\ Q_{ec} \end{bmatrix} = \begin{bmatrix} T_{eced11} & T_{eced12} \\ T_{eced21} & T_{eced22} \end{bmatrix} \begin{bmatrix} P_{ed} \\ Q_{ed} \end{bmatrix}. \tag{A.2}$$

The transfer matrices for the middle ear and outer ear can be multiplied together to give a single expression for the pressure and volume velocity at the ear canal in terms of those at the stapes:

$$\begin{bmatrix} P_{ec} \\ Q_{ec} \end{bmatrix} = \begin{bmatrix} T_{ecst11} & T_{ecst12} \\ T_{ecst21} & T_{ecst22} \end{bmatrix} \begin{bmatrix} P_{st} \\ Q_{st} \end{bmatrix}. \tag{A.3}$$

A.1.1 Calculating input impedances in a two-port network

In order to calculate the input impedance or admittance at a given terminal of the two-port network, it is necessary to load the other terminal of the two-port network with an impedance. This impedance is either added at the output terminal, Z_{out}, or at the input terminal, Z_{in}.

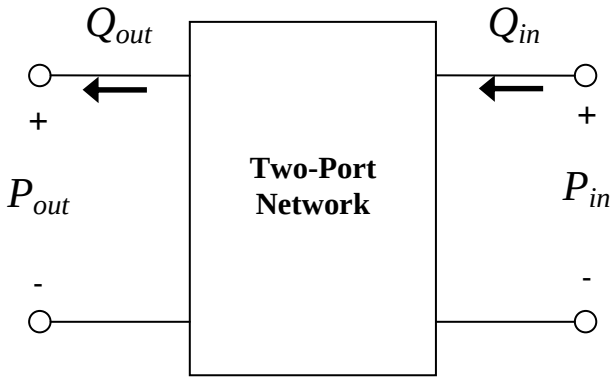


Figure A.3: Two-port network representation given a load termination at the output port.

The general two-port transmission network is given by,

$$\begin{bmatrix} P_{out} \\ Q_{out} \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \begin{bmatrix} P_{in} \\ Q_{in} \end{bmatrix}. \tag{A.4}$$

When the expression for P_{out} is divided by that for Q_{out}, the result is an expression that relates the impedance seen at the output terminal, Z_{out}, given a loading impedance at the input terminal, Z_{in}:

$$Z_{out} = \frac{T_{11}Z_{in} + T_{12}}{T_{21}Z_{in} + T_{22}}, \tag{A.5}$$

where

$$Z_{in} = \frac{P_{in}}{Q_{in}}. \tag{A.6}$$

In order to produce the expressions at the input terminal as a function of those at the output terminal, the inverse transmission matrix must be calculated:

$$\begin{bmatrix} P_{in} \\ Q_{in} \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}^{-1} \begin{bmatrix} P_{out} \\ Q_{out} \end{bmatrix}, \tag{A.7}$$

where

$$\begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}^{-1} = \begin{bmatrix} T_{22} & T_{12} \\ T_{21} & T_{11} \end{bmatrix} \Big/ \left(T_{11} * T_{22} - T_{21} * T_{12} \right). \tag{A.8}$$

If the transmission matrix is reciprocal, satisfying

$$T_{edst11} * T_{edst22} - T_{edst21} * T_{edst12} = 1, \tag{A.9}$$

then the impedance seen at the input terminal, Z_{in}, given a load impedance at the output terminal, Z_{out}, can be written as:

$$Z_{in} = \frac{T_{22} Z_{out} + T_{12}}{T_{21} Z_{out} + T_{11}}. \tag{A.10}$$

A.1.2 Independent responses of the ear canal and middle ear models

Ear canal model

The ear canal is modelled as a hollow cylinder with rigid walls, closed at the eardrum and open at the pinna, unless otherwise noted. Figure A.4 illustrates the model, and Table A.1 presents the physical quantities of the model. Note that the effective length of the ear canal may shorten when an earplug-shielded probe is inserted into the cavity.

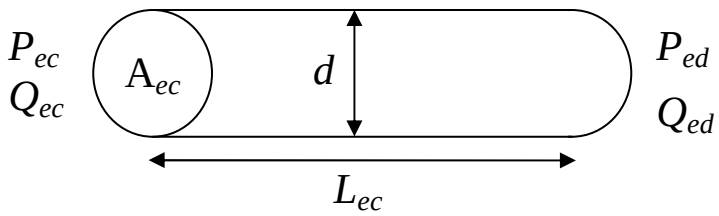


Figure A.4: Schematic illustration of the ear canal model.

A_{ec} is the cross-sectional area of the ear canal, d is its diameter, and L_{ec} is its length.

Quantity	Formula [SI units]
A_{ec}	$3.85*10^{-5}$ [m^2]
d	0.007 [m]
L_{ec}	0.02 [m]

Table A.1: Physical characteristics of the ear canal.

It is assumed in the model that propagation of sound is planar. However, above a cut-off frequency, f_c, the model is no longer valid as radial modes begin to affect the response. This value is approximately given by

$$f_c = 0.586 \, c_0 / 2d \,, \tag{A.11}$$

where c_0 is the speed of sound in air (Kinsler, 1982). The cut-off frequency is approximately 29 kHz, so the plane wave assumption can be considered valid across the frequency range of interest, 20Hz – 20 kHz.

The two-port formulation that relates inputs at the eardrum to outputs at the external opening of the ear canal can be visualised as in Figure A.5, and can be given by rewriting equation (A.2)**:**

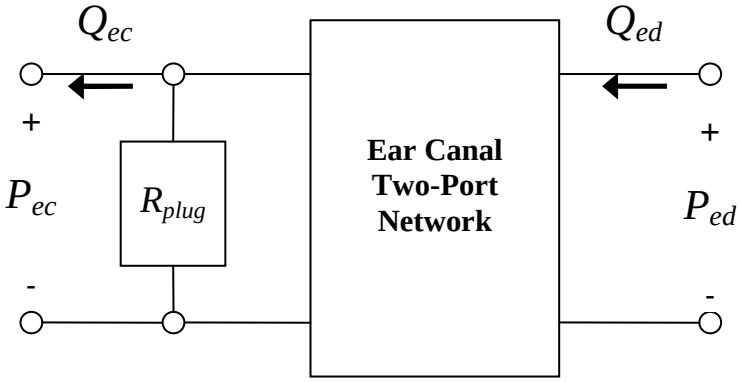


Figure A.5: Two-port network representation of the ear canal.

$$\begin{bmatrix} P_{ec} \\ Q_{ec} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1/R_{plug} & 1 \end{bmatrix} \begin{bmatrix} T_{eced11} & T_{eced12} \\ T_{eced21} & T_{eced22} \end{bmatrix} \begin{bmatrix} P_{ed} \\ Q_{ed} \end{bmatrix}, \tag{A.12}$$

where R_{plug} represents the loss due to the foam plug of an OAE probe at the ear canal; this value is set to $2.2*10^7$ Acoustic Ohms, which is approximately twice the characteristic impedance of air. In practice, this boundary modification reduces the sharpness of the ear

canal resonances. The general form for input impedance in a pipe is given by equation (9.4) in Kinsler (1982). This is reproduced here, but in terms of acoustic impedances (rather than specific acoustic impedances, as given in the original text):

$$\frac{Z_{load}}{\rho_0 c_0 / A} = \frac{\dfrac{Z_{in}}{\rho_0 c_0 / A} + j\tan(kL)}{j\dfrac{Z_{in}}{\rho_0 c_0 / A} + 1}, \tag{A.13}$$

which can be written as

$$Z_{load} = \frac{Z_{in}\cos(kL) + j\sin(kL)\rho_0 c_0 / A}{Z_{in} j\sin(kL) A/\rho_0 c_0 + \cos(kL)}. \tag{A.14}$$

Thus, the elements of the transfer matrix for the outer ear are found to be

$$\begin{array}{ll} T_{eced11} = \cos(kL_{ec}) & T_{eced12} = j*\sin(kL_{ec})\rho_0 c_0 / A_{ec} \\ T_{eced21} = j*\sin(kL_{ec}) A_{ec} / \rho_0 c_0 & T_{eced22} = \cos(kL_{ec}) \end{array}. \tag{A.15}$$

The use of a complex propagation constant accounts for the absorption of sound in the thermal and viscous boundary layers in pipes:

$$k = \frac{\omega}{c_0} - j\alpha, \tag{A.16}$$

given

$$\alpha = 2.89\times10^{-5}\frac{\sqrt{f}}{d/2}, \tag{A.17}$$

where f is the frequency of the driving tone, and d is the diameter of the pipe (Kinsler, 1982).

Figure A.6 shows the input admittance of the ear canal, as seen from the eardrum given a variety of boundary conditions at the external opening of the ear canal.

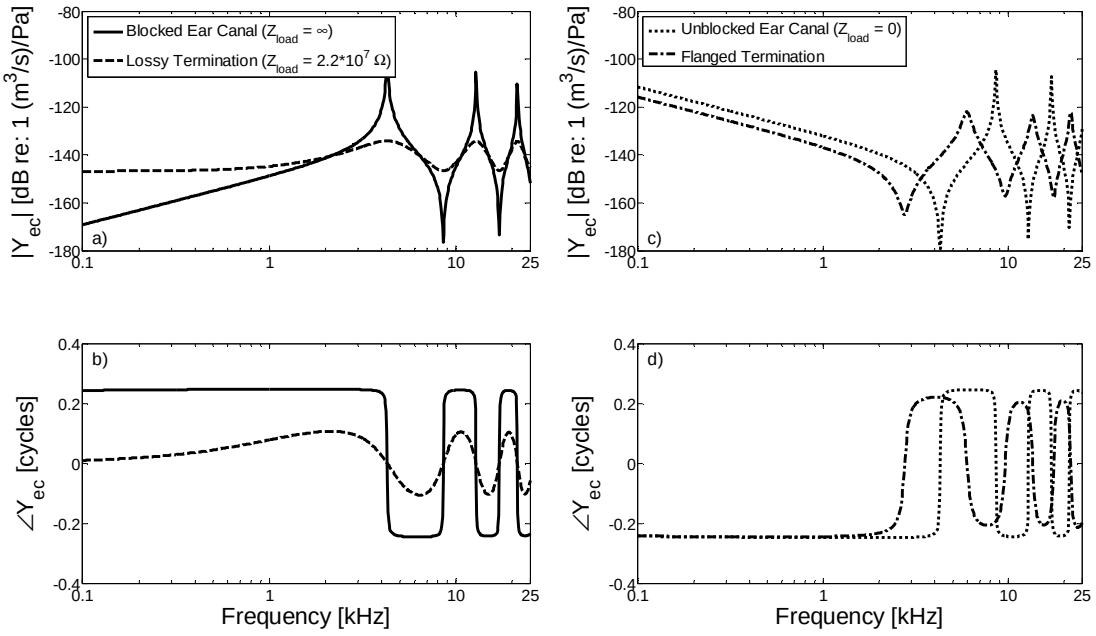


Figure A.6.a-d: Ear canal admittance as seen from the eardrum given blocked (solid), lossy (dashed), flanged (dash-dot) and unblocked (dotted) terminations.

The blocked condition at the end of the ear canal simulates the situation when an OAE is measured with a hard earplug and probe that block the ear canal. Note that this admittance is small at low frequencies and its resonances which correspond to wavelengths of $\lambda = 2*L$, $2/3*L$, etc. where L is the length of the ear canal. Conversely, the resonances for the open-ended system are given by wavelengths of $\lambda = 4*L, 4/3*L$, etc.

When OAEs are measured in practice, the ear is often sealed with a foam earplug which surrounds the receiver and transmitter. In order to simulate this condition, a lossy load can be incorporated at the end of the ear, as in equation (A.12). Similarly, the sharpness of the unblocked resonances is also reduced when the flanged boundary is applied, similar to a real ear canal that opens out into the pinna. In contrast to the lossy boundary, the flanged termination results in a phase shift due to its imaginary component.

Middle ear model

The middle ear consists of three bones: the malleus, incus, and stapes, as seen in Figure A.7. As mentioned previously, the primary function of the middle-ear is to improve the impedance matching between the air in the ear canal and the fluid in the cochlea. This is accomplished through a reduction in the surface area from the tympanic membrane to

the stapes, and also through the mechanical advantage of the lever arm of the malleus and incus. This results in an increase in efficiency of the transfer of energy to the inner ear.

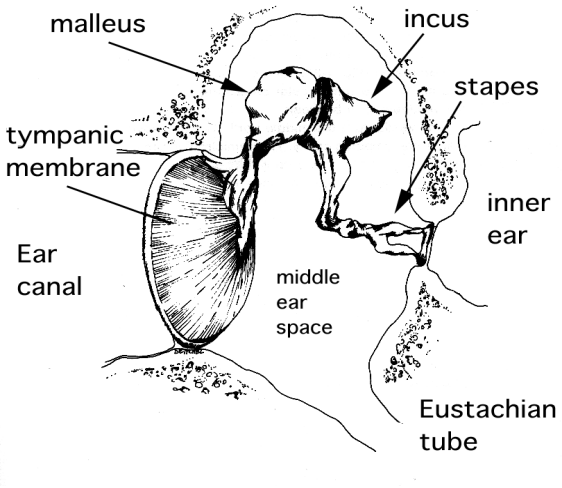


Figure A.7: Human middle ear. Reproduced with permission (Coleman, 2008).

There are many middle ear models that have been reported on in the literature, for example O'Conner and Puria (2008), Pascal (1998), Kringlebotn (1988) and Zwislocki (1962). The model chosen here was proposed by Kringlebotn (1988), and can be expressed in a two-port network formulation, which is illustrated in Figure A.8, and defined in equation (A.1):

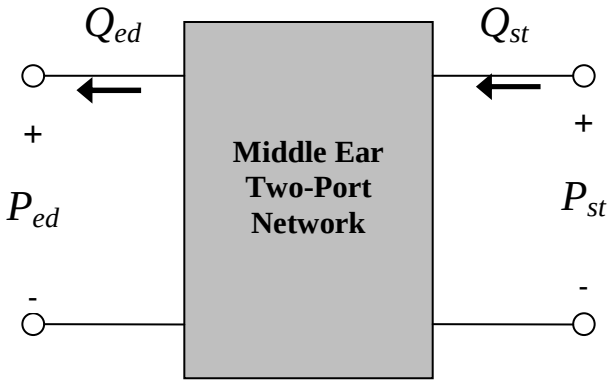


Figure A.8: Two-port network representation of the middle ear.

$$\begin{bmatrix} P_{ed} \\ Q_{ed} \end{bmatrix} = \begin{bmatrix} T_{edst11} & T_{edst12} \\ T_{edst21} & T_{edst22} \end{bmatrix} \begin{bmatrix} P_{st} \\ Q_{st} \end{bmatrix}. \tag{A.1}$$

Figure A.9 and Table A.2 illustrate and identify the internal impedances of the model, which includes separate models of the eardrum (tympanic membrane), middle ear bones

(malleus, incus and stapes), and the stapes footplate. These are coupled together by transformers representing the changes in area or geometry, as described below.

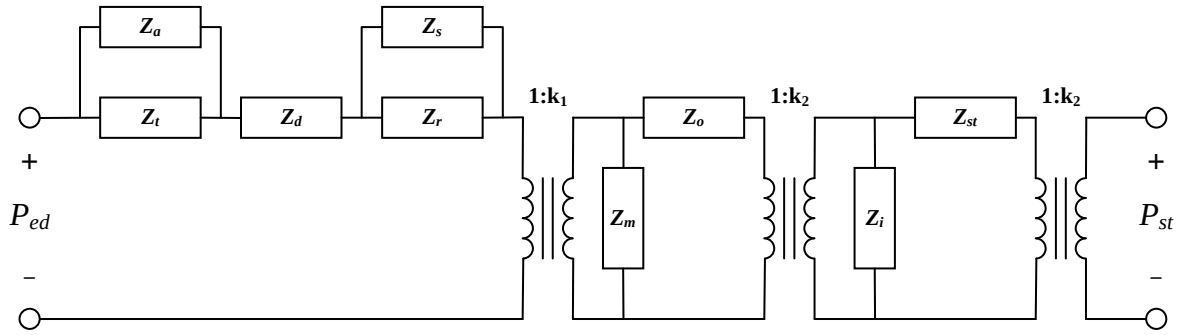


Figure A.9: Diagram of the middle ear impedances in the model of Kringlebotn (1988).

The transformer ratios shown in Figure A.9 correspond to the anatomical area of the eardrum (k_1), the ratio of the effective lever arm of the malleus divided by the lever arm of the incus (k_2), and the inverse of the area of the stapes footplate (k_3).

Expression for Impedance	**Mechanical Quantity being Modelled**
$Z_a = R_a + i\omega L_a + 1/(i\omega C_a)$	antrum and mastoid cells
$Z_t = 1/(i\omega C_t)$	tympanic cavity
$Z_d = i\omega Ld$	drum
$Z_s = R_s + i\omega L_s + 1/(i\omega C_s)$	suspension of the eardrum
$Z_r = R_r + 1/(i\omega C_r)$	rim of the eardrum
$Z_m = R_m + 1/(i\omega C_m)$	coupling between the malleus and incus
$Z_o = R_o + i\omega L_o + 1/(i\omega C_o)$	ossicles (malleus and incus)
$Z_i = R_i + 1/(i\omega C_i)$	coupling between the incus and stapes
$Z_{st} = i\omega L_{st} + 1/(i\omega C_{st})$	stapes, stapedius tendon, oval window

Table A.2: List of middle ear impedances and physical quantities being modelled.

It should be noted that the stapes impedance, Z_{st}, is comprised of the inertial and compliance terms listed as L_c and C_c in the Kringlebotn paper. R_c is the impedance of the cochlea in Kringlebotn's paper, but is omitted in this two-port model as the goal is to terminate this middle ear model with an impedance representing the Neely and Kim (1986) model of the cochlea.

The numerical quantities for the inertial, compliance, resistance terms, and transformer ratios are given in Table A.3 (original cgs units) and Table A.4 (SI units):

Inertial terms [dyn*s^2/cm^5]	**Compliance terms [cm^5/dyn]**	**Resistance terms [dyn*s/cm^5]**
$L_a = 1*10^{-3}$	$C_a = 3.9*10^{-6}$	$R_a = 60$
	$C_t = 0.4*10^{-6}$	
$L_d = 7.5*10^{-3}$		
$L_s = 66*10^{-3}$	$C_s = 0.3*10^{-6}$	$R_s = 20$
	$C_r = 1.3*10^{-6}$	$R_r = 120$
	$C_m = 0.38*10^{-6}$	$R_m = 120$
$L_o = 22*10^{-3}$	$C_o = inf$	$R_o = 200$
	$C_i = 0.3*10^{-6}$	$R_i = 6000$
$L_{st} = 46*10^{-3}$	$C_{st} = 0.56*10^{-6}$	
Transformer Ratios		
$k_1 = 0.6$ [cm^2]	$k_2 = 1.3$ [unitless]	$k_3 = 31.25$ [cm^{-2}]

Table A.3: Parameters of the middle ear model in cgs units as given in Kringlebotn (1988).

Inertial terms [N*s^2/m^5]	**Compliance terms [m^5/N]**	**Resistance terms [N*s/m^5]**
$L_a = 1*10^{2}$	$C_a = 3.9*10^{-11}$	$R_a = 6*10^{6}$
	$C_t = 4*10^{-12}$	
$L_d = 7.5*10^{2}$		
$L_s = 6.6*10^{3}$	$C_s = 3*10^{-12}$	$R_s = 2*10^{6}$
	$C_r = 1.3*10^{-11}$	$R_r = 1.2*10^{7}$
	$C_m = 3.8*10^{-12}$	$R_m = 1.2*10^{7}$
$L_o = 2.2*10^{3}$	$C_o = inf$	$R_o = 2*10^{7}$
	$C_i = 5.6*10^{-12}$	$R_i = 6*10^{8}$
$L_{st} = 4.6*10^{3}$	$C_{st} = 5.6*10^{-12}$	
Transformer Ratios		
$k_1 = 6*10^{-5}$ [m^2]	$k_2 = 1.3$ [unitless]	$k_3 = 3.125*10^{5}$ [m^{-2}]

Table A.4: Parameters of the middle ear model in SI units as converted from Kringlebotn (1988).

Note that the value of R_o has been increased by a factor of 10 to better match measured results.

The network model of Figure A.10 can be simplified by combining the parallel and series impedances at the eardrum into a single term Z_1. Similarly, Z_m, Z_o, Z_i and Z_{st} are denoted Z_2, Z_3, Z_4, and Z_5, respectively:

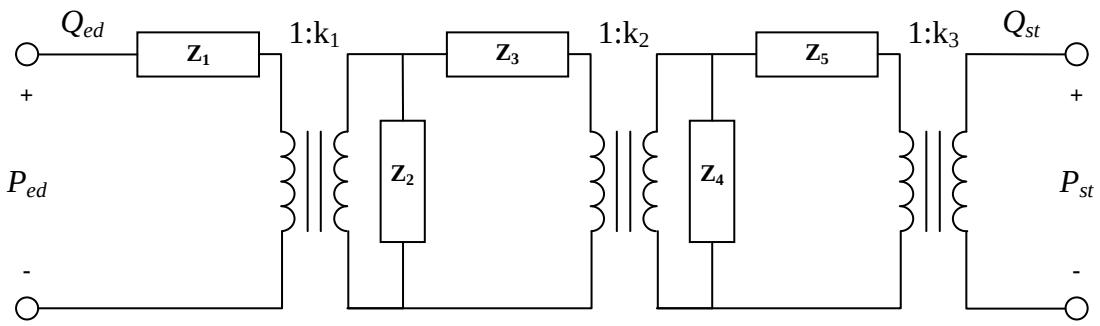


Figure A.10 Simplified block diagram of the network model of the middle ear.

A two port network model of the middle-ear can be derived from the following relationships:

$$\begin{aligned} T_{edst11} &= \left.\frac{P_{ed}}{P_{st}}\right|_{Q_{st}=0} & T_{edst12} &= \left.\frac{P_{ed}}{Q_{st}}\right|_{P_{st}=0} \\ T_{edst21} &= \left.\frac{Q_{ed}}{P_{st}}\right|_{Q_{st}=0} & T_{edst22} &= \left.\frac{Q_{ed}}{Q_{st}}\right|_{P_{st}=0} \end{aligned} . \qquad \text{(A.18)}$$

The transmission matrix elements expressed in terms of the middle ear impedances and transformer ratios are thus determined to be:

$$\begin{aligned} T_{edst11} &= \left[\frac{(Z_1+Z_2)(Z_2+Z_3+Z_4)}{Z_2Z_4} - \frac{Z_2}{Z_4}\right] \Big/ k_1k_2k_3 & T_{edst12} &= \left[\frac{(Z_1Z_2+Z_1Z_3+Z_2Z_3)(Z_4+Z_5)+(Z_1+Z_2)(Z_4Z_5)}{Z_2Z_4}\right] * k_1k_2k_3 \\ T_{edst21} &= \left[\frac{(Z_2+Z_3+Z_4)}{Z_2Z_4}\right] \Big/ k_1k_2k_3 & T_{edst22} &= \left[\frac{(Z_2+Z_3)(Z_4+Z_5)+Z_4Z_5}{Z_2Z_4}\right] * k_1k_2k_3 \end{aligned} . \qquad \text{(A.19)}$$

It is possible to show this is a reciprocal network, since

$$T_{edst11} * T_{edst22} - T_{edst21} *_{Tedst12} = 1 . \qquad \text{(A.20)}$$

The calculated admittance of the middle ear two-port network as viewed from the stapes when the eardrum is unblocked and blocked is shown in Figure A.11.a-b.

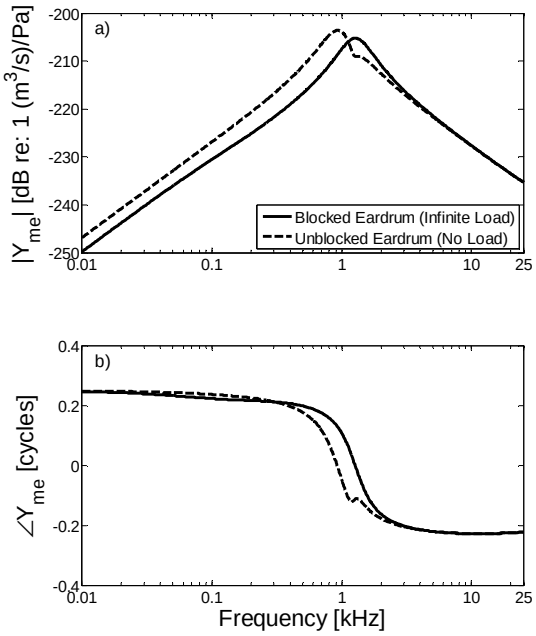


Figure A.11.a-b: Input admittance of the middle ear as seen from the stapes given blocked (solid) and unblocked (dashed) conditions at the eardrum.

A.2 Response of the combined middle ear and ear canal models

The coupled response of the middle ear and ear canal two-port networks can be combined, as shown in Figure A.2, to give the overall transmission matrix shown in Figure A.12:

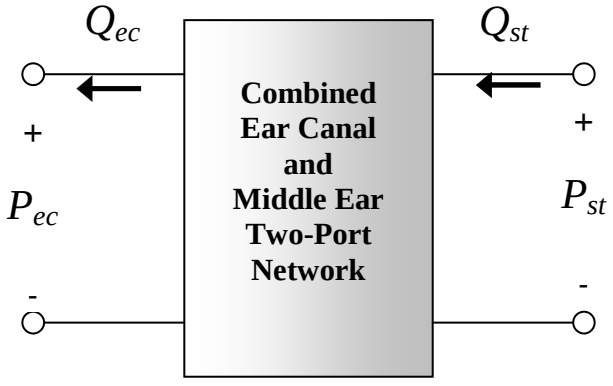


Figure A.12: Two-port network representation of the combined middle ear and ear canal.

This result has the transmission matrix shown in equation (A.21):

$$\begin{bmatrix} P_{ec} \\ Q_{ec} \end{bmatrix} = \begin{bmatrix} T_{ecst11} & T_{ecst12} \\ T_{ecst21} & T_{ecst22} \end{bmatrix} \begin{bmatrix} P_{st} \\ Q_{st} \end{bmatrix}. \tag{A.21}$$

A.2.1 Forward transfer function in frequency

In order to generate an appropriate stimulus at the stapes, it is important to consider both the input at the ear canal, and the input to the state space model. Previous work done by Harte (2004) showed that the loading on a typical in-ear loudspeaker due to the ear canal space is negligible. As a result, the volume displacement of the probe can be considered proportional to its input voltage. However, the input to the state space cochlea is a linear acceleration, as the model is formulated in specific acoustic impedances. Thus, the desired transfer function is the ratio of stapes acceleration to volume displacement in the ear canal.

The two-port network is formulated in terms of volume velocities. The forward volume velocity transfer function can be determined by expanding the second row of equation (A.3). This gives

$$Q_{ec} = T_{ecst21} P_{st} + T_{ecst22} Q_{st}. \tag{A.22}$$

Dividing both sides of this equation by Q_{st} and inverting becomes

$$\frac{Q_{st}}{Q_{ec}} = \left(T_{ecst21} Z_{st} + T_{ecst22} \right)^{-1}, \tag{A.23}$$

where Z_{st} is the input impedance of the cochlea. For the purposes of Figure A.13, Z_{st} is taken as $1.1*10^{10}$ acoustic ohms. Equation (A.23) is related to the required transfer function in the following manner:

$$\frac{a_{st}}{\int Q_{ec}} = \frac{s^2}{A_{st}} * \frac{Q_{st}}{Q_{ec}}, \tag{A.24}$$

where a_{st} is the acceleration of the stapes footplate, the symbol $\int Q_{ec}$ represents the volume displacement in the ear canal, A_{st} is the area of the stapes footplate and s is equal to $j\omega$. Figure A.14 illustrates this transfer function when the ear canal is terminated with the lossy impedance representing the foam earplug.

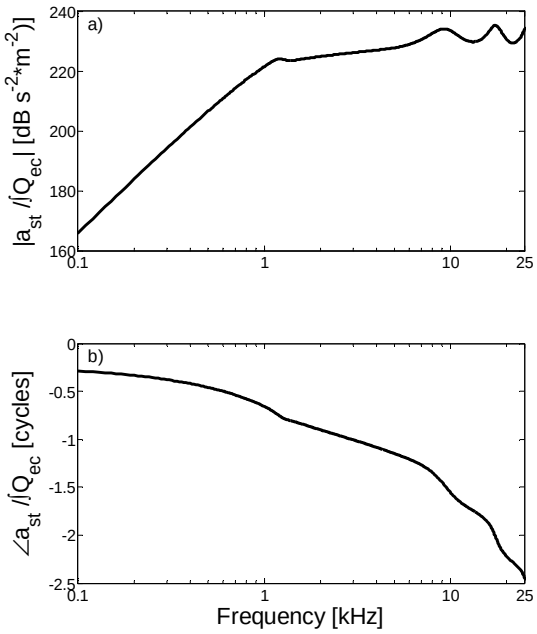


Figure A.13: Magnitude (a) and phase (b) of the forward transfer function frequency response: stapes acceleration given a volume displacement in the ear canal. The impedance of the cochlea is taken to be $1.1*10^{10}$ Acoustic Ohms.

A.2.2 Forward transfer function in time

In order to generate the stapes acceleration due to a time-varying volume displacement in the ear canal, it is necessary to convolve the input volume displacement with the impulse response of the transfer function:

$$a_{st}(t) = \left[\int^{t} Q_{ec}(t')\,dt' \right] \otimes h_{a_{st}/\int Q_{ec}}(t) \tag{A.25}$$

Figure A.14 illustrates the acceleration at the stapes due to a 100 μs pulse in volume displacement at the ear canal. The magnitude of the volume displacement pulse was chosen to produce the same RMS stapes acceleration as a 40 dB peak SPL pulse of pressure of 100 μs duration presented at the eardrum. This is referred to as a 'standard click' in this work. Although the magnitude of the stapes acceleration decays quite rapidly, there are still some low-amplitude oscillations that ring on for some time.

When the amplitude of the click is greatly increased, the low-level residual oscillations of the stimulus can affect the response in the cochlea. In order to force the click response to zero after 10 ms without introducing a sudden change in stimulus level, the decreasing half of a 10 ms Hanning window is applied to the input after 5 ms have

passed. This is only important at very high levels, as the magnitude of the response after 5 ms is less than 1% of the amplitude at the peak.

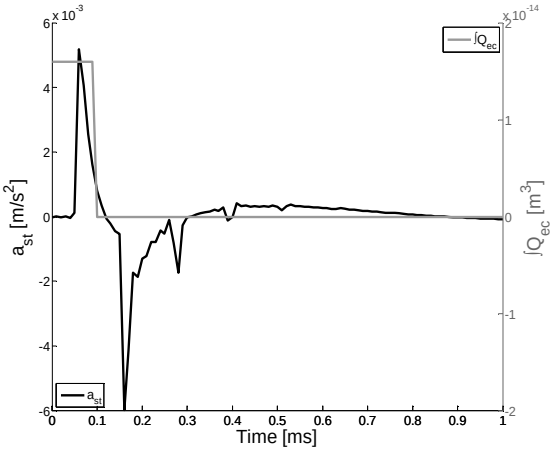


Figure A.14: Standard click input to the model: stapes acceleration (black, left vertical axis) due to an ear canal volume displacement (gray, right vertical axis).

A.2.3 Reverse transfer function in frequency

The reverse transfer function is defined to be the ratio of an output ear canal pressure due to an input stapes pressure. In order to calculate this quantity, it is necessary to take the inverse of the transfer matrix that relates the pressure and volume velocity. Recall the general solution of a two-by-two matrix when the direction of the volume velocity is accounted for:

$$\begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}^{-1} = \begin{bmatrix} T_{22} & T_{12} \\ T_{21} & T_{11} \end{bmatrix} \Big/ \left(T_{11} * T_{22} - T_{21} * T_{12} \right). \tag{A.8}$$

As the transfer matrix in equation (4.3) is reciprocal, the forward transfer matrix relationship is thus given as

$$\begin{bmatrix} P_{st} \\ Q_{st} \end{bmatrix} = \begin{bmatrix} T_{ecst22} & T_{ecst12} \\ T_{ecst21} & T_{ecst11} \end{bmatrix} \begin{bmatrix} P_{ec} \\ Q_{ec} \end{bmatrix}, \tag{A.26}$$

where the elements of the matrix are given by equation (A.8). By expanding the first line, we have

$$P_{st} = T_{ecst22} P_{ec} + T_{ecst12} Q_{ec}. \tag{A.27}$$

Similarly to the previous transfer function calculation, both sides of the equation are divided by P_{ec} and inverted to give

$$\frac{P_{ec}}{P_{st}} = \left(T_{ecst22} + T_{ecst12} \Big/ Z_{ec} \right)^{-1}, \tag{A.28}$$

where Z_{ec} is the impedance seen at the ear canal. As the lossy nature of the ear canal boundary is already taken into account with the matrix in (A.2), Z_{ec} is simply set to infinity. The reverse pressure transfer function is shown in Figure A.15.

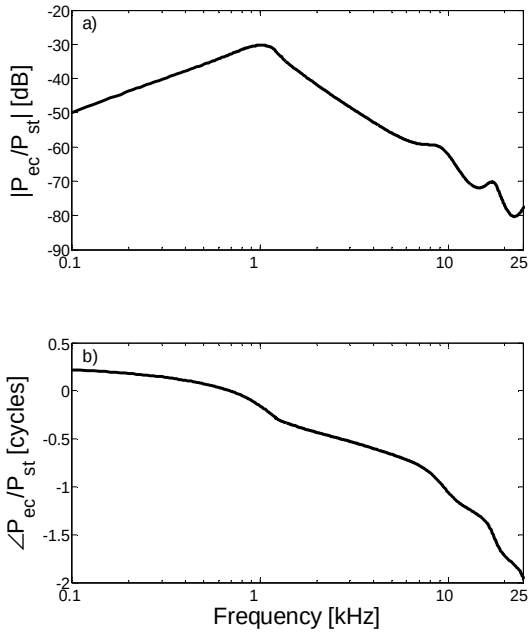


Figure A.15: Reverse pressure frequency response: pressure at the outer ear per unit input pressure at the stapes.

A.2.4 Reverse transfer function in time

The pressure in the ear canal due to a time-varying pressure at the stapes can be determined by convolving the stapes pressure with the impulse response of the reverse-pressure transfer function:

$$P_{ec}(t) = P_{st}(t) \otimes h_{P_{ec}/P_{st}}(t). \tag{A.29}$$

The reverse-pressure response given by a 100 μs pulse of pressure at the stapes of unit amplitude, as calculated by the convolution of equation (A.29), is shown in Figure A.16.

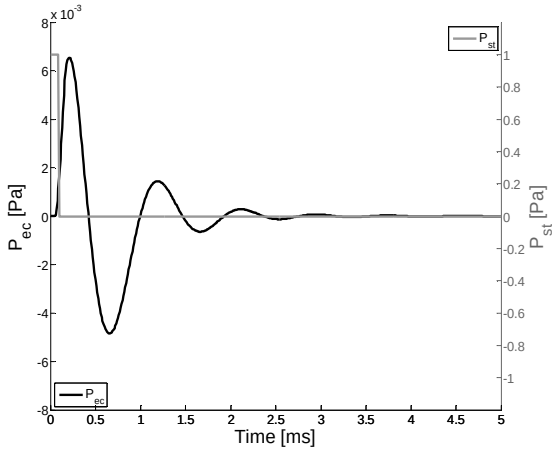


Figure A.16: Reverse pressure transfer function: pressure at the outer ear per unit input pressure at the stapes.

A.3 Validation of the two-port model

There are a number of sources against which the two-port model of the middle and outer ears can be validated. The most important of these is a comparison with measured physiological data. The results from Puria (2003) are chosen for this purpose as measurements were taken with two-port modelling and consequences for OAEs in mind. Figure A.17.a-b compares forward and backward pressure gain through the two-port network with measured results, where

$$\mathrm{M1} = \frac{P_{st}}{P_{ec}}, \tag{A.30}$$

and

$$\mathrm{M2} = \frac{P_{ec}}{P_{st}}. \tag{A.31}$$

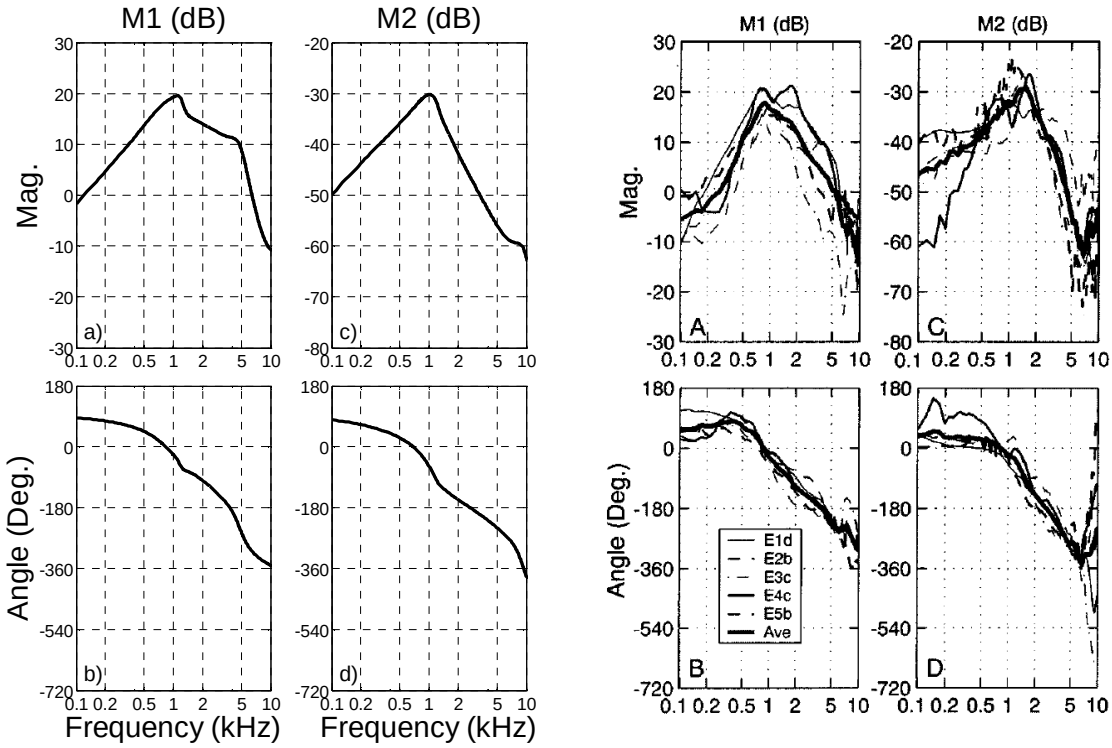


Figure A.17.a-b: Comparison of forward and reverse pressure gain functions: (a-d) two-port results for forward pressure gain, P_{st}/P_{ec} (M1), and reverse pressure gain, P_{ec}/P_{st} (M2); and (A-D) five experimentally measured curves (Puria, 2003). A value of $Z_{st} = 1.1*10^{10}$ acoustic ohms was used to generate the two-port (a) forward pressure gain (M1).

Figure A.18.a-b illustrates the product of M1 and M2, thus generating the round-trip gain through the middle ear,

$$G_{ME}^{RT} = \mathrm{M1} \cdot \mathrm{M2} \,. \tag{A.32}$$

It can be seen that all of the predicted results from the two-port model are in reasonably good agreement with Puria's (2003) measurements, both in terms of the magnitudes and the phases.

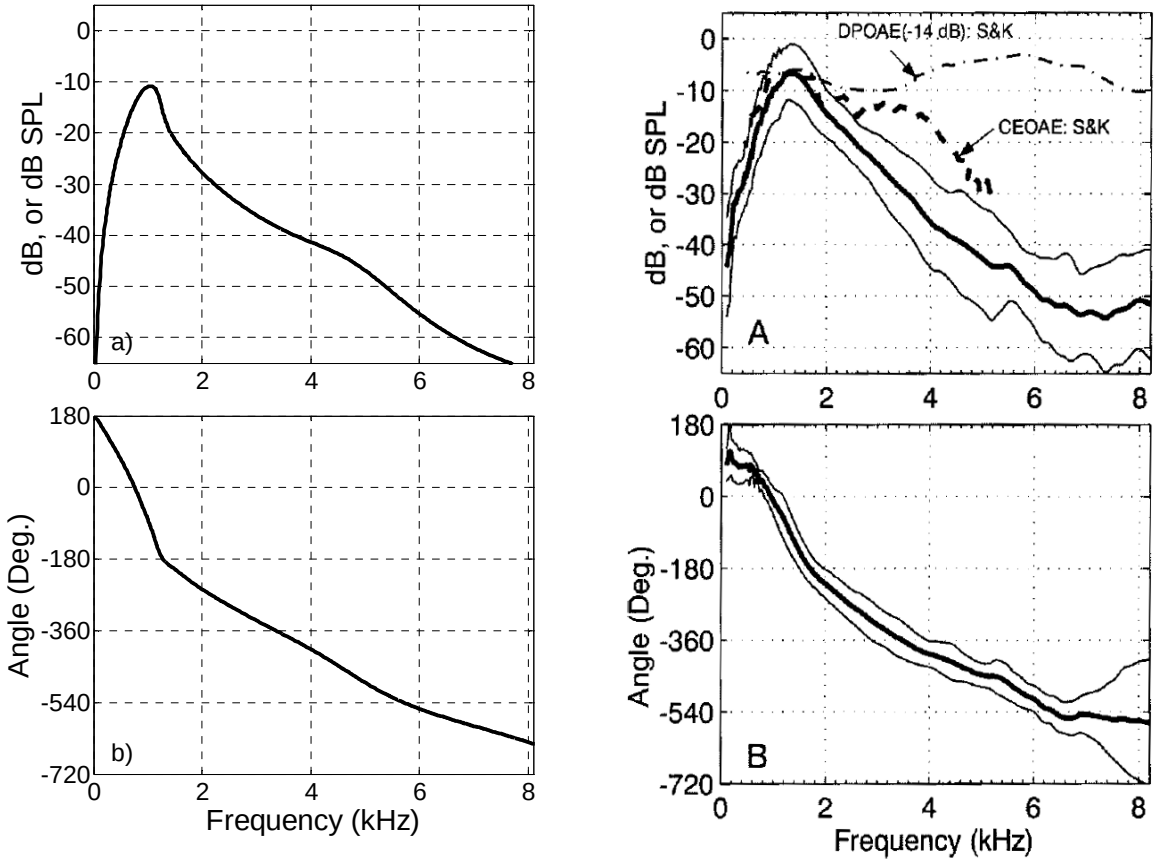


Figure A.18.a-b: Comparison of round-trip middle ear gain as produced by the two-port model (a, b) and as experimentally measured (A, B) by Puria (2003). The dotted lines in (A) represent click- and distortion product-otoacoustic emission levels measured by Smurzynski and Kim (1992).

A.4 Reverse input admittance in state space

In order to integrate the frequency domain network models of the middle ear and ear canal into the existing state space model of the cochlea, the boundary condition at the stapes must be reformulated in state space. While there are straightforward methods of converting transfer function representations of physical systems into state space, such as that of the middle ear, there is no obvious analytic method of converting the existing two-port model of the ear canal into state space. The solution has been to simply fit the reverse middle ear impedance (Z_{out}) with a mass-spring-damper system, where

$$Z_{out} = \frac{P_{ed}}{Q_{st}}. \tag{A.33}$$

The model parameters are listed below in Table A.5 alongside the parameters derived by Puria (2003).

	Model parameters	Puria (2003)
Mass (mg)	2.96	1.41
Damping (N*s/m^3)	2.79*10^4	3.20*10^4
Stiffness (N/m^3)	2.63*10^8	2.59*10^8

Table A.5: Lumped element parameters of the middle ear boundary.

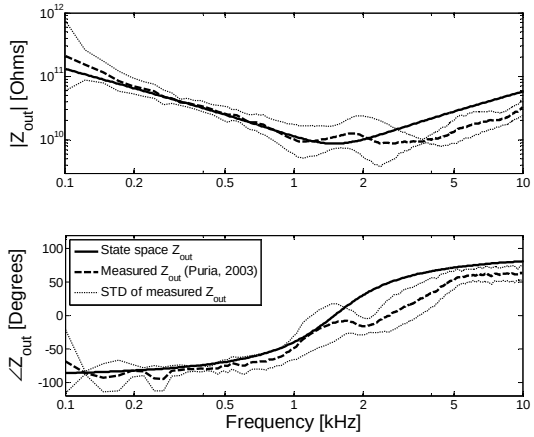


Figure A.19.a-b: Comparison of model and measured reverse middle ear impedance frequency response (Puria, 2003).

A.5 Summary and Discussion

The combined transfer functions of the two-port network agree quite well with measured results. Puria's external measurements were made in the ear canal so presumably the pressure microphone affected the apparent impedance of the ear canal. The two-port results are close enough to average human measurements that they should provide a valid approximation of a true middle-ear and ear canal response given physiological variations exist between individuals. In addition, the two-port network's round-trip middle ear gain agrees qualitatively with the equivalent physiologically-based plot.

In summary, the forward- and reverse- transfer functions can be applied either in time or in frequency to relate measurements and stimuli in the ear canal to outputs and inputs in the cochlear model, respectively. The only potential downside of this implementation is that the input impedance of the cochlea, Z_c, is assumed to be constant. Puria's (2003) experimental measurements show that this quantity varies less than tenfold within frequencies of 0.1 to 10 kHz in human cadavers. However, this may change in a situation where there are cochlear reflections that propagate back to the stapes. This may

represent a source of experimental error if one attempts to study stimulus-frequency OAEs with this model. Nevertheless, the input impedance of the cochlea should not change appreciably during the brief input applied by a click stimulus, for instance. Thus, this implementation of the middle and outer ears can be considered valid for the investigations presented in this thesis.

Appendix B

Derivation of Neely and Kim Quantities

In this appendix, a variety of quantities pertaining to the Neely and Kim (1986) formulation are derived. In section B.1, the ratio of BM and TM micromechanical motion to a driving pressure is determined. Section B.2 presents the undamped natural frequencies of the two micromechanical masses. Section B.3 shows how the stability criterion of an isolated, active micromechanical element can found. Finally, Section B.4 arranges the equations of motion for the micromechanical model in a form suitable for the state space model.

B.1 BM and TM admittance

The impedances of the micromechanical model as defined by Neely & Kim (1986) are:

$$\begin{aligned} Z_1 &= k_1/s + c_1 + sm_1 \\ Z_2 &= k_2/s + c_2 + sm_2 \\ Z_3 &= k_3/s + c_3 \\ Z_4 &= k_4/s + c_4 \end{aligned}, \tag{B.1}$$

where

$$s = j\omega, \tag{B.2}$$

and Z_1 is the impedance associated with the organ of corti, Z_2 the TM, Z_3 the coupling between the organ of corti and the TM, and Z_4 provides 'a frequency-dependent phase-shift

between Pa and ξ_c' (Neely and Kim, 1986). ξ_c, the difference in the motion of the BM and the TM is given by

$$\xi_c(x) = g(x)\xi_b(x) - \xi_t(x). \tag{B.3}$$

$P_a(x)$ is the active pressure generated by the OHCs:

$$P_a(x) = -\gamma Z_4(x)\dot{\xi}_c(x). \tag{B.4}$$

The equations of motion for the BM is

$$P_d(x,t) - P_a(x,t) = gZ_1(x)\dot{\xi}_b(x,t) + Z_3(x)\dot{\xi}_c(x,t), \tag{B.5}$$

and the motion of the TM can be described by:

$$0 = Z_2(x)\dot{\xi}_t(x,t) - Z_3\dot{\xi}_c(x,t). \tag{B.6}$$

We start by taking (A.5) and replacing $\dot{\xi}_t(x,t)$ with a substitution from (A.2):

$$\begin{aligned} \xi_t(x,t) &= g(x)\xi_b(x,t) - \xi_c(x,t) \\ \frac{d}{dt}(\xi_t(x,t)) &= \dot{\xi}_t(x,t) = g(x)\dot{\xi}_b(x,t) - \dot{\xi}_c(x,t) \end{aligned} \tag{B.7}$$

$$\begin{aligned} 0 &= Z_2(x)\left(g(x)\dot{\xi}_b(x,t) - \dot{\xi}_c(x,t)\right) - Z_3\dot{\xi}_c(t,x) \\ 0 &= Z_2(x)g(x)\dot{\xi}_b(x,t) - (Z_2 + Z_3)\dot{\xi}_c(t,x) \\ \dot{\xi}_c(t,x) &= \frac{Z_2(x)g(x)\dot{\xi}_b(x,t)}{(Z_2 + Z_3)} \end{aligned} \tag{B.8}$$

and then substitute solution from (B.7) into (B.3):

$$P_a(x) = -\gamma Z_4(x)\frac{Z_2(x)g(x)\dot{\xi}_b(x,t)}{(Z_2 + Z_3)}. \tag{B.9}$$

Next substitute solution from (B.7) into (B.4):

$$P_d(x,t) - P_a(x,t) = gZ_1(x)\dot{\xi}_b(x,t) + Z_3(x)\frac{Z_2(x)g(x)\dot{\xi}_b(x,t)}{(Z_2 + Z_3)}. \tag{B.10}$$

Then substitute (B.9) into (B.10) and solve for the BM velocity:

$$\begin{aligned} P_d(x,t) - \left(-\gamma Z_4(x)\frac{Z_2(x)g(x)\dot{\xi}_b(x,t)}{(Z_2 + Z_3)}\right) &= g(x)Z_1(x)\dot{\xi}_b(x,t) + Z_3(x)\frac{Z_2(x)g(x)\dot{\xi}_b(x,t)}{(Z_2 + Z_3)} \\ P_d(x,t) &= \dot{\xi}_b(x,t)\left[g(x)Z_1(x) + Z_3(x)\frac{Z_2(x)g(x)}{(Z_2(x) + Z_3(x))} - \gamma Z_4(x)\frac{Z_2(x)g(x)}{(Z_2(x) + Z_3(x))}\right] \\ \Rightarrow \dot{\xi}_b(x,t) &= P_d(x,t) \Big/ g(x)\left[Z_1(x) + Z_2(x)\left(\frac{Z_3(x) - \gamma Z_4(x)}{Z_2(x) + Z_3(x)}\right)\right] \end{aligned} \tag{B.11}$$

This result agrees with equation (13) in Neely & Kim's paper (1986), where the denominator of (A.10) is described as the 'driving-point impedance of the CP' in equation

(12) which includes an added term b that is cancelled out in (13). $\dot{\xi}_b(x,t)$ is integrated by dividing by s on both sides to return $\xi_b(x,t)$:

$$\xi_b(x,t) = P_d(x,t) \Big/ s * g(x)\left[Z_1(x) + Z_2(x)\left(\frac{Z_3(x) - \gamma Z_4(x)}{Z_2(x) + Z_3(x)} \right) \right] \tag{B.12}$$

The transfer function of the BM receptance is given by $\frac{\xi_b(x,t)}{P_d(x,t)}$:

$$\frac{\xi_b(x,t)}{P_d(x,t)} = \frac{1}{s * g(x)\left[Z_1(x) + Z_2(x)\left(\frac{Z_3(x) - \gamma Z_4(x)}{Z_2(x) + Z_3(x)} \right) \right]}, \tag{B.13}$$

which can be rewritten as

$$\frac{\xi_b(x,t)}{P_d(x,t)} = \frac{Z_2(x) + Z_3(x)}{s * g(x)\left[Z_1(x)\left(Z_2(x) + Z_3(x) \right) + Z_2(x)\left(Z_3(x) - \gamma Z_4(x) \right) \right]}. \tag{B.14}$$

The frequency response function of the TM can be expressed as a function of the BM frequency response function. Summing (B.5) and (B.6) gives

$$P_d(x,t) - P_a(x,t) = gZ_1(x)\dot{\xi}_b(x,t) + Z_2(x)\dot{\xi}_t(x,t). \tag{B.15}$$

Substitute in equation (B.8):

$$\begin{aligned} P_d(x,t) - \left[-\gamma Z_4(x)\left(\dot{\xi}_b - \dot{\xi}_t \right) \right] &= gZ_1(x)\dot{\xi}_b(x,t) + Z_2(x)\dot{\xi}_t(x,t) \\ \dot{\xi}_t(x,t)\left[Z_2(x) + \gamma Z_4(x) \right] &= P_d(x,t) - \dot{\xi}_b(x,t)\left[gZ_1(x) - \gamma Z_4(x) \right] \end{aligned}. \tag{B.16}$$

Thus, our solution for the TM admittance frequency response function is:

$$\frac{\dot{\xi}_t(x,t)}{P_d(x,t)} = \frac{1 - \left[\dot{\xi}_b(x,t) / P_d(x,t) \right]\left[gZ_1(x) - \gamma Z_4(x) \right]}{\left[Z_2(x) + \gamma Z_4(x) \right]}. \tag{B.17}$$

B.2 Undamped natural frequencies

In order to better understand this model's micromechanical response as a function of position along the BM, it is useful to investigate the undamped natural frequencies of the system. To find resonances, set $p_a(t)=c_1=c_2=c_3=0$ and evaluate the equations of motion:

$$\begin{aligned} 0 &= -M_1\ddot{x}_1 - K_1x_1 + K_3(x_2 - x_1) \\ 0 &= M_1\ddot{x}_1 + x_1(K_1 + K_3) - x_2K_3 \end{aligned}, \tag{B.18}$$

and

$$\begin{aligned} 0 &= M_2\ddot{x}_2 + K_2x_2 + K_3(x_2 - x_1) \\ 0 &= M_2\ddot{x}_2 + x_2(K_2 + K_3) - x_1K_3 \end{aligned}, \tag{B.19}$$

Set (1) and (2) in matrix form:

$$\begin{bmatrix} M_1 & 0 \\ 0 & M_2 \end{bmatrix}\begin{bmatrix} \ddot{x}_1 \\ \ddot{x}_2 \end{bmatrix} + \begin{bmatrix} K_1 + K_3 & -K_3 \\ -K_3 & K_2 + K_3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{B.20}$$

Substitute sinusoidal solutions for x_1 and x_2 back into (B.19) and (B.20):

$$\begin{aligned} x_1 &= A_1\sin(\omega t + \alpha) \\ x_2 &= A_2\sin(\omega t + \alpha) \end{aligned} \tag{B.21}$$

which yields

$$\begin{aligned} &-\omega^2A_1M_1\sin(\omega t + \alpha) + A_1\sin(\omega t + \alpha)*(K_1 + K_3) - A_2\sin(\omega t + \alpha)K_3 = 0 \\ &\Rightarrow A_1(-\omega^2M_1 + K_1 + K_3) - A_2K_3 = 0 \end{aligned} \tag{B.22}$$

and

$$\begin{aligned} &-\omega^2A_2M_2\sin(\omega t + \alpha) + A_2\sin(\omega t + \alpha)*(K_2 + K_3) - A_1\sin(\omega t + \alpha)K_3 = 0 \\ &\Rightarrow A_2(-\omega^2M_2 + K_2 + K_3) - A_1K_3 = 0 \end{aligned}. \tag{B.23}$$

The determinant of the coefficients of A_1 and A_2 in equation (2.9) is called the characteristic determinant:

$$\Delta(\omega) = \begin{vmatrix} K_1 + K_3 - \omega^2M_1 & -K_3 \\ -K_3 & K_2 + K_3 - \omega^2M_2 \end{vmatrix} = 0 \tag{B.24}$$

.

Equating the characteristic determinant with zero gives the frequency equation:

$$\begin{aligned} &\omega^4(M_1M_2) - \omega^2(K_1M_2 + K_3M_2 + K_2M_1 + K_3M_1) + (K_1K_2 + K_1K_3 + K_2K_3) = 0 \\ &\Rightarrow \omega^4 - \omega^2\left(\frac{K_1 + K_3}{M_1} + \frac{K_2 + K_3}{M_2}\right) + \frac{(K_1K_2 + K_1K_3 + K_2K_3)}{M_1M_2} = 0 \end{aligned} \tag{B.25}$$

.

Solving the frequency equation in terms of ω^2 gives the following result:

$$\omega^2 = \left(\frac{K_1 + K_3}{2M_1} + \frac{K_2 + K_3}{2M_2}\right) \pm \sqrt{\left(\frac{K_1 + K_3}{2M_1} + \frac{K_2 + K_3}{2M_2}\right)^2 - \frac{(K_1K_2 + K_1K_3 + K_2K_3)}{M_1M_2}} \tag{B.26}$$

.

This returns two positive solutions for ω^2, which correspond to the solutions for the resonant frequencies, ω_1 and ω_2. This is applied in Chapter 2.

B.3 Frequency domain stability criterion

The stability of a frequency domain model that contains active feedback can be determined by isolating its system dynamics from its feedback function.

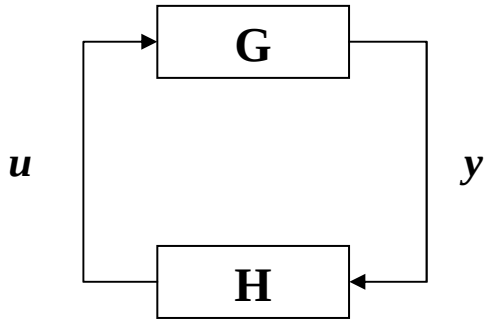


Figure B.1: General diagram of micromechanical controller.

The quantities involved in the solution for the system G and H are given in Table B.2:

$u = y * H$	(B.27)	$u = \dot{\xi}_c(x) * [-\gamma Z_4(x)] = P_a(x)$	(B.30)
$G = \frac{y}{u}$	(B.28)	$G = \frac{\dot{\xi}_c(x)}{P_a(x)}$	(B.31)
$y = g\dot{\xi}_b(x) - \dot{\xi}_t(x) = \dot{\xi}_c(x)$	(B.29)	$H = -\gamma Z_4(x)$	(B.32)

Table B.2: quantities involved in determining the system plant *G*, and feedback block, *H*.

ξ_c is the difference in position between the BM and the TM:

$$\xi_c(x) = g(x)\xi_b(x) - \xi_t(x) \tag{B.33}$$

$P_a(x)$ is the active pressure generated by the OHCs.

$$P_a(x) = -\gamma Z_4(x)\dot{\xi}_c(x) \tag{B.34}$$

Combine (B.33) and (B.34):

$$P_a(x) = -\gamma Z_4(x)\left[g(x)\dot{\xi}_b(x) - \dot{\xi}_t(x)\right] \tag{B.35}$$

Equations of motion are as follows:

$$P_d(t,x) - P_a(t,x) = gZ_1(x)\dot{\xi}_b(t,x) + Z_3(x)\dot{\xi}_c(t,x), \tag{B.36}$$

and for the second degree of freedom:

$$0 = Z_2(x)\dot{\xi}_t(x) - Z_3\dot{\xi}_c(x). \tag{B.37}$$

The goal is to solve these equations to find the transfer function G, $\frac{\dot{\xi}_c(x)}{P_a(x)}$. Begin by solving (B.37) for $\dot{\xi}_t(x)$:

$$\dot{\xi}_t(x) = \frac{Z_3}{Z_2}\dot{\xi}_c(x). \tag{B.38}$$

Recall from (B.33) that

$$\dot{\xi}_b(x) = \frac{\dot{\xi}_c(x) + \dot{\xi}_t(x)}{g(x)}. \tag{B.39}$$

Substitute (B.38) into (B.39):

$$\dot{\xi}_b(x) = \frac{\dot{\xi}_c(x) + \frac{Z_3}{Z_2}\dot{\xi}_c(x)}{g(x)} = \frac{\dot{\xi}_c(x)}{g(x)}\left(1 + \frac{Z_3}{Z_2}\right). \tag{B.40}$$

Set the input $P_d = 0$ and rearrange terms in the first degree of motion (B.36) with a substitution from (B.40).

$$\begin{aligned} 0 - P_a(t,x) &= gZ_1(x)\left[\frac{\dot{\xi}_c(x)}{g(x)}\left(1 + \frac{Z_3}{Z_2}\right)\right] + Z_3(x)\dot{\xi}_c(t,x) \\ -P_a(t,x) &= \dot{\xi}_c(x)\left[Z_1(x)\left(1 + \frac{Z_3}{Z_2}\right) + Z_3(x)\right]. \end{aligned} \tag{B.41}$$

Now expand terms:

$$-P_a(t,x) = \dot{\xi}_c(x)\left(\frac{Z_1Z_2 + Z_1Z_3 + Z_2Z_3}{Z_2}\right). \tag{B.42}$$

Solve for G, recalling (B.31):

$$G = \frac{\dot{\xi}_c(x)}{P_a(x)} = -\left(\frac{Z_2}{Z_1Z_2 + Z_1Z_3 + Z_2Z_3}\right). \tag{B.43}$$

Multiply the solution for *G* by *H* to obtain *GH*:

$$GH = \gamma Z_4\left(\frac{Z_2}{Z_1Z_2 + Z_1Z_3 + Z_2Z_3}\right). \tag{B.44}$$

B.4 BM and TM equations of motion

The equations of motion for the BM and TM in Neely and Kim's (1986) micromechanical model can be derived by summing the forces acting on each mass. Figure 2.1 illustrates Neely and Kim's micromechanical model.

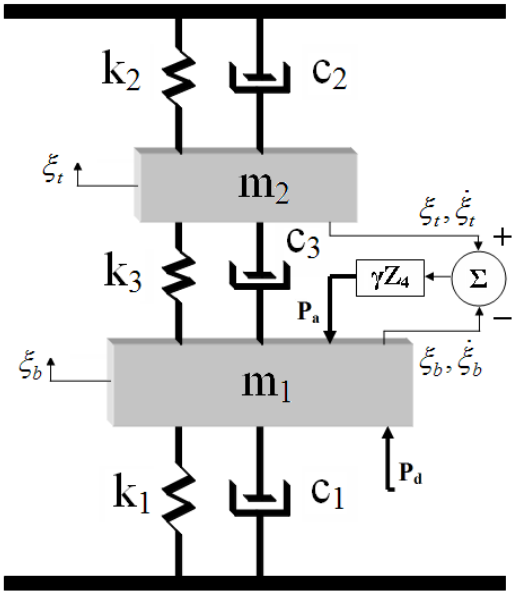


Figure 2.1: Neely & Kim's (1986) micromechanical model of the cochlea.

The forces acting on each mass are summarized in Figure B.2.

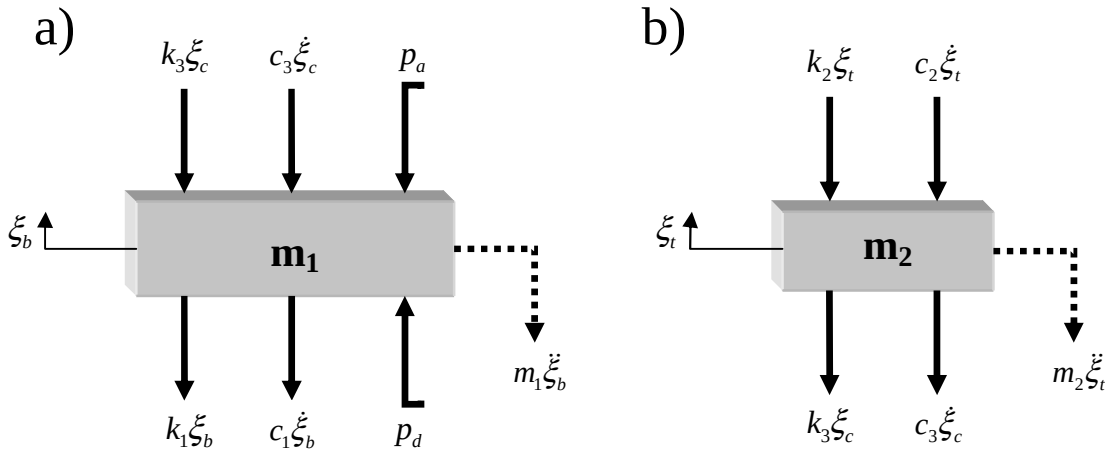


Figure B.2.a-b: Force diagram for the BM (a) and the TM (b) in Neely and Kim's (1986) micromechanical model. A dotted line represents the inertial forces due to the accelerations of the masses.

The equations of motion for the BM and the TM are thus

$$p_d - p_a = \left[m_1 \ddot{\xi}_b + c_1 \dot{\xi}_b + k_1 \xi_b \right] + \left[c_3 \dot{\xi}_c + k_3 \xi_c \right], \tag{B.45}$$

and

$$0 = \left[m_2 \ddot{\xi}_t + c_2 \dot{\xi}_t + k_2 \xi_t \right] - \left[c_3 \dot{\xi}_c + k_3 \xi_c \right], \tag{B.46}$$

respectively, where $\xi_c = \xi_b - \xi_t$, and $p_a = -\gamma \left[c_4 \dot{\xi}_c + k_4 \xi_c \right]$. Solving equations (B.45) and (B.46) in terms of the accelerations of each mass returns the format required for a state space formulation:

$$\ddot{\xi}_b = \frac{1}{m_1} \left\{ p_d + \gamma \left[c_4 \dot{\xi}_c + k_4 \xi_c \right] - \dot{\xi}_b \left(c_1 + c_3 \right) - \xi_b \left(k_1 + k_3 \right) + \dot{\xi}_t c_3 + \xi_t k_3 \right\}, \tag{B.47}$$

and

$$\ddot{\xi}_t = \frac{1}{m_2} \left\{ -\dot{\xi}_t \left(c_2 + c_3 \right) - \xi_t \left(k_2 + k_3 \right) + \dot{\xi}_b c_3 + \xi_b k_3 \right\}. \tag{B.48}$$

Appendix C

Journal Publications

C.1 Elliott et al. *2007*

This journal paper concerns the formulation of a state space representation for cochlear mechanics, and specifically applied to the Neely and Kim (1986) model of the cochlea. The text was largely written by S.J. Elliott, but all code, associated simulations and figures were developed by E.M. Ku.

Appendix C

A state space model for cochlear mechanics

Stephen J. Elliott,[a] Emery M. Ku, and Ben Lineton
Institute of Sound and Vibration Research, University of Southampton, Southampton, Hampshire SO17 1BJ, United Kingdom

(Received 14 March 2007; revised 14 August 2007; accepted 14 August 2007)

The stability of a linear model of the active cochlea is difficult to determine from its calculated frequency response alone. A state space model of the cochlea is presented, which includes a discretized set of general micromechanical elements coupled via the cochlear fluid. The stability of this time domain model can be easily determined in the linear case, and the same framework used to simulate the time domain response of nonlinear models. Examples of stable and unstable behavior are illustrated using the active micromechanical model of Neely and Kim. The stability of this active cochlea is extremely sensitive to abrupt spatial inhomogeneities, while smoother inhomogeneities are less likely to cause instability. The model is a convenient tool for investigating the presence of instabilities due to random spatial inhomogeneities. The number of unstable poles is found to rise sharply with the relative amplitude of the inhomogeneities up to a few percent, but to be significantly reduced if the spatial variation is smoothed. In a saturating nonlinear model, such instabilities generate limit cycles that are thought to produce spontaneous otoacoustic emissions. An illustrative time domain simulation is presented, which shows how an unstable model evolves into a limit cycle, distributed along the cochlea. © *2007 Acoustical Society of America.*
[DOI: 10.1121/1.2783125]

PACS number(s): 43.64.Kc, 43.64.Jb, 43.40.Vn, 43.64.Bt [BLM] Pages: 2759–2771

I. INTRODUCTION

It is important to develop mathematical models of cochlear mechanics to test our understanding of the physical processes involved. Modeling accomplishes this by providing predictions of complex system responses, based on assumptions regarding the underlying physical processes, which can be compared with direct measurements. Linear models of the cochlea, such as those presented by Zwislocki (1950), Allen (1977), and Neely and Kim (1986), as reviewed by de Boer (1996) for example, can provide predictions of the distribution of motion along the cochlea at a given frequency, or of its frequency response at a given position. These models include the macromechanical behavior of the fluid coupling along the length of the cochlea, as well as the micromechanical behavior of the individual parts of the cochlear partition. Such models are the starting point for more realistic nonlinear models. In order to produce numerical results, the cochlear partition, which has mechanical parameters that vary continuously along its length, is often approximated by a discrete set of elements. This allows a finite dimensional set of equations to be solved one frequency at a time (Neely, 1981). The number of elements is generally quite large, typically about 500, so that a single model generates many hundreds of individual frequency response functions.

One disadvantage of such a frequency domain formulation is that it can be difficult to determine whether the system being modeled is stable or not. With an active micromechanical element to represent the cochlear amplifier, energy can be supplied to the system at any point along the cochlea and so the stability of the overall coupled model is not guaranteed.

The stability of cochlear models is important for several reasons. For an unstable system, a frequency domain model will still yield results. However, these results must be treated with some care, since the assumption of a fixed amplitude output for a sinusoidal input has been violated. Also, it may not always be obvious that the model is unstable simply from inspection of the predicted frequency response, possibly leading to the misinterpretation of these results. This is particularly true of cochlear models containing distributed inhomogeneities (e.g., Zweig and Shera, 1995), where the predicted responses involve both anterograde and retrograde traveling waves. In the current paper, this problem is overcome by the development of a state space model that allows a rigorous approach to the study of model stability.

A second benefit of this model is that it facilitates the study of spontaneous otoacoustic emissions (SOAEs). For example, with this approach, it is easy to determine unambiguously whether or not a given cochlear model will lead to the generation of SOAEs; this in turn enables the examination of the conditions that lead to SOAE generation.

Two aspects of the stability of cochlear models that have been previously discussed in the literature are the stability of a single isolated point on the basilar membrane, and the stability of the entire cochlear model after all points on the basilar membrane have been coupled together via fluid interactions. Kanis and de Boer (1993, p. 3201) suggest that a single isolated point on the basilar membrane will be unstable if the resistance (i.e., the real part of the driving point impedance) is negative at any frequency. However, while a negative resistance is a necessary condition for instability, it is not sufficient. As pointed out by Zweig (1991, p. 1246), a

[a]Electronic mail: sje@isvr.soton.ac.uk

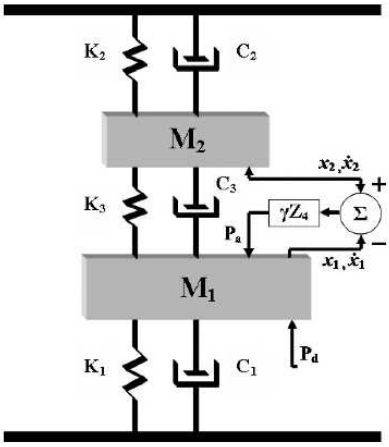


FIG. 1. Equivalent mechanical system for Neely and Kim's active micromechanical model.

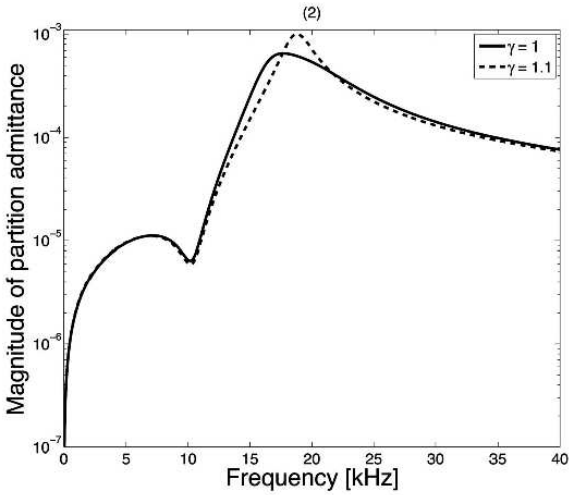


FIG. 2. Magnitude of the partition admittance of an isolated micromechanical element for feedback gains of γ=1 and 1.1.

system may have a negative resistance and still remain stable, provided the real parts of all its poles are negative.

While it is relatively straightforward to determine the stability of a single uncoupled point on the basilar membrane, this task is much more difficult when dealing with the full cochlear model. As for the single point, the suggestion by Kanis and de Boer (1993), and van Hengel (comment in Shera and Zweig, 1993, p. 62) that instability will arise whenever the resistance of the driving point is negative, is not necessarily true. Koshigoe and Tubis (1983) suggested using the Hilbert transform to check that causality (and hence stability) was implied by the resulting frequency responses of the model. While this is a valid method of determining instability, it has the disadvantage that responses over a wide range of frequencies must be obtained, and that (at least in theory) the frequency response of every elemental point in the model must be checked.

As an example of the frequency domain analysis of an unstable system, we consider the calculated response of a single, isolated, active micromechanical element as described by Neely and Kim (1986). Their micromechanical element is a two degree of freedom system, as illustrated in Fig. 1, in which an active pressure, p_a, acts on the lower mass, which is proportional to the relative displacement of the two masses. The admittance of this isolated element of the cochlear partition can be written as

$$Y_p(j\omega) = (b/g)[Z_1(j\omega) + Z_2(j\omega)(Z_3(j\omega) - \gamma Z_4(j\omega))/(Z_2(j\omega) + Z_3(j\omega)]^{-1}, \quad (1)$$

in which

$$Z_1(j\omega) = \frac{K_1}{j\omega} + C_1 + j\omega M_1, \quad Z_2(j\omega) = \frac{K_2}{j\omega} + C_2 + j\omega M_2,$$

$$Z_3(j\omega) = \frac{K_3}{j\omega} + C_3$$

and

$$Z_4(j\omega) = \frac{K_4}{j\omega} + C_4,$$

where b is the ratio of the maximum to average basilar membrane (BM) displacement, g is the lever gain, the masses and stiffnesses are all defined in Fig. 1, and γ is the gain of the feedback loop used to model the cochlear amplifier.

Equation (1) corresponds to the frequency response of the velocity of M_1 with respect to the pressure acting upon it, and its magnitude is plotted in Fig. 2 for the parameters listed in Table I at x=5 mm, for feedback gains of γ=1 and γ=1.1. Although the calculated frequency response for γ=1.1 is smoothly varying and apparently plausible, it has no physical significance in isolation, since the system is un-

TABLE I. Parameters of the micromechanical model of Neely and Kim (1986) with corrections for original typographical errors and converted into SI units.

Quantity	Formula (SI)
$K_1(x)$	$1.1^{*}10^{10}e^{-400x}$ N*m^{-3}
$C_1(x)$	$200+15\,000e^{-200x}$ N*s*m^{-3}
M_1	$3^{*}10^{-2}$ kg*m^{-2}
$K_2(x)$	$7^{*}10^{7}e^{-440x}$ N*m^{-3}
$C_2(x)$	$100e^{-220x}$ N*s*m^{-3}
M_2	$5^{*}10^{-3}$ kg*m^{-2}
$K_3(x)$	$1^{*}10^{8}e^{-400x}$ N*m^{-3}
$C_3(x)$	$100e^{-80x}$ N*s*m^{-3}
$K_4(x)$	$6.15^{*}10^{8}e^{-400x}$ N*m^{-3}
$C_4(x)$	$100e^{-80x}$ N*s*m^{-3}
gamma	1
g	1
b	0.4
L	0.025 m
K_s	$2.1^{*}10^{6}$ N*m^{-3}
C_s	4000 N*s*m^{-3}
M_s	$45^{*}10^{-2}$ kg*m^{-2}
ρ	1000 kg*m^{-3}

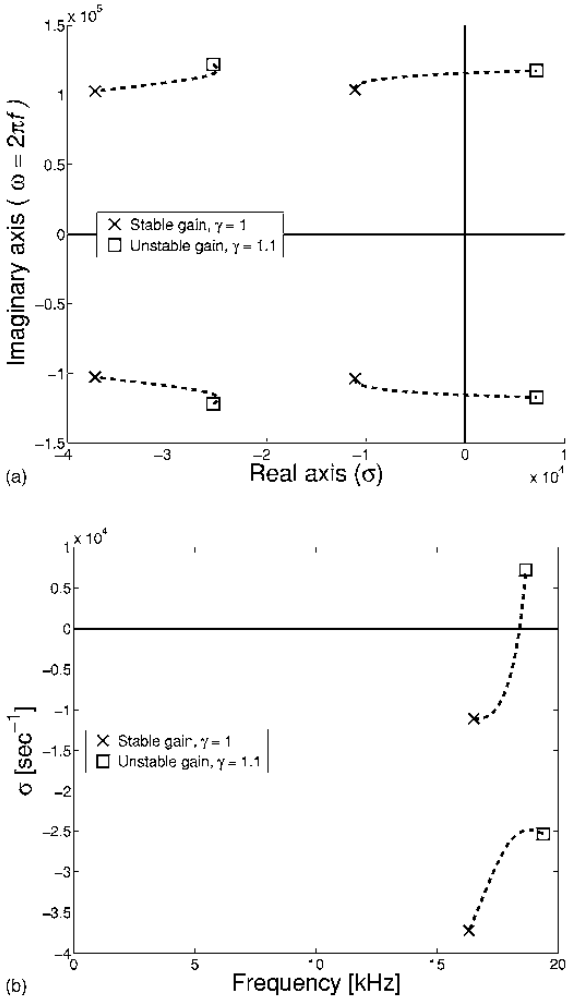


FIG. 3. Positions of the poles of the transfer function for the micromechanical element with feedback gains of $\gamma=1$ and $\gamma=1.1$, plotted: (a) using the conventional control representation with positive and negative frequencies plotted vertically; and (b) using the representation used for the cochlea model described here, with frequency plotted horizontally.

stable for this value of feedback gain and so the response to any sinusoidal input would be exponentially diverging, and ultimately infinite.

The stability of this single input, single output system can be determined by writing its transfer function as

$$Y_p(s) = (b/g)\{Z_1(s) + Z_2(s)[Z_3(s) - \gamma Z_4(s)]/ [Z_2(s) + Z_3(s)]\}^{-1}, \quad (2)$$

and calculating the values of the complex Laplace transform variable, s, for which $Y_p(s)$ tends to infinity, i.e., the poles of the system. Since the system has two degrees of freedom, it has two pairs of poles, which are plotted in the complex s plane for the system with feedback gains of $\gamma=1$ and $\gamma=1.1$ in Fig. 3. The positions of these poles are plotted in two ways in Figs. 3(a) and 3(b). Figure 3(a) shows the conventional representation used by the control community, with the real part of s (σ) along the horizontal axis and both the negative and positive components of the imaginary part of s ($\omega=2\pi f$) along the horizontal axis. The poles appear in complex conjugate pairs, so no additional information is gained by plotting those with a negative imaginary part. To help emphasise the physical significance of the poles, their imaginary parts can be divided by 2π, to give units of Hz, and the frequency plotted along the horizontal axis, as in a conventional frequency response, which is shown in Fig. 3(b).

One such pair of poles would give rise to a transient response proportional to $e^{\sigma t}\sin(2\pi f t)$. The real part of the pole position, σ, which is plotted on the vertical axis in Fig. 3(b), can thus be physically interpreted as the divergence rate of the transient. If σ is negative, the transient converges and the system is stable, but if σ becomes positive, the transient diverges and the system is unstable.

The two poles with positive real parts that appear when $\gamma=1.1$ thus clearly indicate that this system is unstable (Franklin *et al.*, 2005). Only the stability of the isolated micromechanical element is being considered in this example. When such an array of elements is coupled together in the cochlea, it is not generally possible to express the overall transfer function of the complete system analytically and hence determine the position of the poles. Although the phase of the calculated frequency response can give an indication of unstable behavior for this single input single output system, as can the causality of the inverse Fourier transform of the frequency response, it is difficult to formalize such tests for discretized cochlear models producing many hundreds of individual frequency responses. This is of particular concern given that numerical effects, such as truncation in the frequency domain, will also lead to the presence of small noncausal components.

The original motivation for developing a state space model of cochlear dynamics is that its stability can be unambiguously determined by examining the eigenvalues of a single matrix. These eigenvalues correspond to the system's poles, as described by Franklin *et al.* (2005), for example, and can be accurately calculated even for large matrices because of the wide availability of efficient numerical algorithms in software such as MATLAB. In such a state space model the dynamics of the underlying states of the system are expressed in the time domain as a set of coupled first-order differential equations (Furuta *et al.* 1988). Any instability of the system, including that of a hidden mode, is guaranteed to be represented in such a model (Skogestad and Postlethwaite, 1996). Several reliable software packages are now available specifically for solving ordinary differential equations cast in this form and for the direct time domain simulation of state space models. The interpretation of these models is also well established in the control engineering community. State space models have been proposed previously to describe the dynamics of the cochlea by Monderer and Lazar (1988), Stibler *et al.* (1998), and Lindgren and Li (2003). In each of these cases, however, the cochlea model was passive and so the stability of the model was not an issue.

A general framework for the state space description of a discretized cochlea model is set out in Sec. II. In Sec. III, the

features of this model are illustrated using Neely and Kim's (1986) formulation of the cochlea's micromechanical behavior. It is shown that when this model is stable, the frequency responses predicted using the state space model are identical to those calculated using the conventional frequency domain formulation for the parameters given by Neely and Kim.

In Sec. IV, the stability of this model is examined in more detail, particularly for variations of the cochlear amplifier's feedback gain along the length of the cochlea. It is shown that the stability of such a coupled cochlea is extremely sensitive to even small abrupt changes in gain, while the stability of the model is maintained for larger changes in feedback gain that occur more smoothly along the length of the cochlea.

The effect of random spatial inhomogeneities along the length of the cochlea is examined in Sec. V, using simulations of random gain perturbations having different magnitudes and length scales. It is shown that the number of unstable poles rises with the magnitude of the random perturbation, but perturbations of much less than 1% can easily produce many unstable poles. In a completely linear system these unstable poles would cause exponentially divergent oscillations. If the feedback forces in the micromechanical model saturate at a particular level, however, the amplitude of these oscillations will stabilize and the response reaches a limit cycle.

This nonlinear behavior is explored in Sec. VI, in which an example time domain simulation is presented, based on the state space model. The response of the cochlear model is computed when compressive nonlinearities are incorporated into each micromechanical feedback loop. A simulation of such a nonlinear cochlear model, which is unstable when linear, is presented showing that the nonlinearity does indeed limit the amplitude of the unstable behavior, resulting in a limit cycle distributed along the length of the cochlea.

II. GENERAL STATE SPACE FORMULATION

In this section, a state space formulation is developed for model of a discretized cochlea as formulated by Neely (1981), and as elaborated upon by Neely and Kim (1986). Adopting the long wavelength assumption, the starting point is the differential equation describing one-dimensional wave propagation along the cochlea (as discussed, for example, by de Boer, 1996)

$$\frac{\partial^2 p(t)}{\partial x^2} - \frac{2\rho}{H}\ddot{w}(t) = 0, \tag{3}$$

where $p(t)$ is the wave form of the pressure difference across the cochlear partition; $\ddot{w}(t)$ is the radially averaged transverse acceleration of the cochlear partition, both of which are also functions of x, although the dependence is suppressed for notational convenience; ρ is the density of the cochlear fluids; and H is the height of the canal above and below the cochlear partition, which is assumed to be constant. The boundary condition for the wave equation at the basal end can be written as

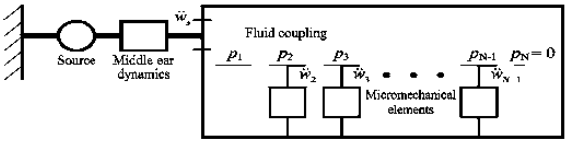


FIG. 4. The discrete model of the cochlea including the micromechanical models of the cochlear partition for elements 2 to $N-1$, the model of the middle ear dynamics at element 1 and of the helicotrema at element N.

$$\left.\frac{\partial p(t)}{\partial x}\right|_{x=0} = 2\rho\ddot{w}_S(t), \tag{4}$$

where $\ddot{w}_S(t)$ is the acceleration of the stapes footplate. This has two linearly superposing components: the unloaded part $\ddot{w}_{SO}(t)$, which is the acceleration due to an external excitation, normally the pressure in the ear canal, and $\ddot{w}_{SR}(t)$, which is the acceleration due to the loading by the internal pressure response in the cochlea at $x=0$. Thus, Eq. (4) may be written as

$$\left.\frac{\partial p(t)}{\partial x}\right|_{x=0} - 2\rho\ddot{w}_{SR}(t) = 2\rho\ddot{w}_{SO}(t). \tag{5}$$

At the apical end of the cochlea, a pressure release boundary condition at the helicotrema can be written as

$$p(t)|_{x=L} = 0. \tag{6}$$

Using finite difference approximations for the spatial derivatives in Eqs. (3), (5), and (6) as originally proposed by Neely (1981), in which the cochlear length L is divided into N sections of length Δ, the wave equation, Eq. (3), can be approximated by

$$\frac{p_{n-1}(t) - 2p_n(t) + p_{n+1}(t)}{\Delta^2} - \frac{2\rho}{H}\ddot{w}_n(t) = 0, \tag{7}$$

for $n=2$ to $n=N-1$, where $p_n(t)$ and $\ddot{w}_n(t)$ are the wave forms of the pressure difference and acceleration of the cochlear partition of the nth element. The basal boundary condition, Eq. (5), can also be approximated by

$$\frac{p_2(t) - p_1(t)}{\Delta} - 2\rho\ddot{w}_{SR}(t) = 2\rho\ddot{w}_{SO}(t). \tag{8}$$

The apical boundary condition, Eq. (6), can be written as

$$p_N(t) = 0. \tag{9}$$

This discrete model is illustrated in Fig. 4. Equations (7)–(9) can be written in matrix form (Neely, 1981), as

$$\frac{H}{2\rho\Delta^2}\begin{bmatrix} -\frac{\Delta}{H} & \frac{\Delta}{H} & & & & & & 0 \\ 1 & -2 & 1 & & & & & \\ 0 & 1 & -2 & 1 & & & & \\ & & \ddots & \ddots & \ddots & & & \\ & & & 1 & -2 & 1 & & 0 \\ & & & & 1 & -2 & 1 & \\ 0 & & & & & 0 & -\frac{2\rho\Delta^2}{H} & \end{bmatrix} \times \begin{bmatrix} p_1(t) \\ p_2(t) \\ \vdots \\ p_{N-1}(t) \\ p_N(t) \end{bmatrix} - \begin{bmatrix} \ddot{w}_{SR}(t) \\ \ddot{w}_2(t) \\ \vdots \\ \ddot{w}_{N-1}(t) \\ 0 \end{bmatrix} = \begin{bmatrix} \ddot{w}_{SO}(t) \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \tag{10}$$

which may also be written as

$$\mathbf{F}\mathbf{p}(t) - \ddot{\mathbf{w}}(t) = \mathbf{q}, \tag{11}$$

where $\mathbf{p}(t)$ and $\ddot{\mathbf{w}}(t)$ are the vectors of pressure differences and cochlear partition accelerations; $\mathbf{F}$ is the finite-difference matrix; and $\mathbf{q}$ is the vector of source terms.

We now cast the dynamics of each micromechanical element on the cochlear partition, for $n=2$ to $N-1$, in the state space form

$$\dot{\mathbf{x}}_n(t) = \mathbf{A}_n\mathbf{x}_n(t) + \mathbf{B}_n p_n(t) \tag{12}$$

$$\dot{w}_n(t) = \mathbf{C}_n\mathbf{x}_n(t) \tag{13}$$

where $\mathbf{x}_n(t)$ is the vector of state variables associated with the internal behavior of the nth micromechanical model.

Similarly, the state space equation describing the loading properties of the middle ear (located in our state space system at $n=1$) can be written as

$$\dot{\mathbf{x}}_1(t) = \mathbf{A}_1\mathbf{x}_1(t) + \mathbf{B}_1 p_1(t) \tag{14}$$

$$\dot{w}_{\mathrm{SR}}(t) = \mathbf{C}_1\mathbf{x}_1(t). \tag{15}$$

Although specific examples are given below for the forms of these micromechanical state space equations, they are kept in their most general form for now to illustrate the structure of the formulation. All of these elemental, uncoupled, state space models can be gathered together in the combined matrix equations

$$\dot{\mathbf{x}}(t) = \mathbf{A}_E\mathbf{x}(t) + \mathbf{B}_E\mathbf{p}(t) \tag{16}$$

$$\dot{\mathbf{w}}(t) = \mathbf{C}_E\mathbf{x}(t), \tag{17}$$

where the vectors are defined as

$$\mathbf{x}^T(t) = [\mathbf{x}_1^T(t)\ \mathbf{x}_2^T(t)\ \cdots\ \mathbf{x}_{N-1}^T(t)\ \mathbf{x}_N^T(t)], \tag{18}$$

$$\dot{\mathbf{w}}^T(t) = [\dot{w}_{\mathrm{SR}}(t)\ \dot{w}_2(t)\ \cdots\ \dot{w}_{N-1}(t)\ 0], \tag{19}$$

$$\mathbf{p}^T(t) = [p_1(t)\ p_2(t)\ \cdots\ p_{N-1}(t)\ p_N(t)], \tag{20}$$

and the matrices, which are block diagonal, are defined as

$$\mathbf{A}_E = \begin{bmatrix} \mathbf{A}_1 & 0 & \cdots & & \\ 0 & \mathbf{A}_2 & & & \\ \vdots & & \ddots & & \vdots \\ & & & \mathbf{A}_{N-1} & 0 \\ & & \cdots & 0 & \mathbf{A}_N \end{bmatrix}, \tag{21}$$

$$\mathbf{B}_E = \begin{bmatrix} \mathbf{B}_1 & 0 & \cdots & & \\ 0 & \mathbf{B}_2 & & & \\ \vdots & & \ddots & & \vdots \\ & & & \mathbf{B}_{N-1} & 0 \\ & & \cdots & 0 & \mathbf{B}_N \end{bmatrix}, \tag{22}$$

$$\mathbf{C}_E = \begin{bmatrix} \mathbf{C}_1 & 0 & \cdots & & \\ 0 & \mathbf{C}_2 & & & \\ \vdots & & \ddots & & \vdots \\ & & & \mathbf{C}_{N-1} & 0 \\ & & \cdots & 0 & \mathbf{C}_N \end{bmatrix}. \tag{23}$$

The overall vector of state variables $\mathbf{x}(t)$ should not be confused with the longitudinal spatial variable x used in Eq. (3).

We now assume that the matrix $\mathbf{F}$ in Eq. (11) is invertible, and express this equation for the fluid coupling within the cochlea in the form

$$\mathbf{p}(t) = \mathbf{F}^{-1}\ddot{\mathbf{w}}(t) + \mathbf{F}^{-1}\mathbf{q}(t). \tag{24}$$

Note from Eq. (17) that $\ddot{\mathbf{w}}(t)$ is equal to $\mathbf{C}_E\dot{\mathbf{x}}(t)$, so that

$$\mathbf{p}(t) = \mathbf{F}^{-1}\mathbf{C}_E\dot{\mathbf{x}}(t) + \mathbf{F}^{-1}\mathbf{q}(t). \tag{25}$$

Substituting this into Eq. (16) allows an overall state space equation for the coupled cochlear with distributed micromechanics and dynamic boundary conditions to be written in the general state space form

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t), \tag{26}$$

where

$$\mathbf{A} = [\mathbf{I} - \mathbf{B}_E\mathbf{F}^{-1}\mathbf{C}_E]^{-1}\mathbf{A}_E, \tag{27}$$

$$\mathbf{B} = [\mathbf{I} - \mathbf{B}_E\mathbf{F}^{-1}\mathbf{C}_E]^{-1}\mathbf{B}_E, \tag{28}$$

$$\mathbf{u} = \mathbf{F}^{-1}\mathbf{q}. \tag{29}$$

The matrix $\mathbf{A}$ is called the system matrix, and it determines the transient response of the system. In particular the eigenvalues of the matrix $\mathbf{A}$ are the poles of the system's transfer function and the real parts of all these eigenvalues must be negative if the transient response is to decay away, so that the system is stable.

III. A SPECIFIC EXAMPLE

In this section we demonstrate how the micromechanical model and boundary conditions used by Neely and Kim (1986) can be set in state space form, and reproduce some illustrative frequency responses from their paper to demon-

strate the equivalence of the state space formulation to the conventional frequency domain formulation for a stable system.

The Neely and Kim (1986) micromechanical model is an active second-order system, which must have two state variables associated with each degree of freedom. The four state variables are chosen to be the displacements of the two masses in Fig. 1, $x_1(t)$ and $x_2(t)$, and their velocities, $\dot{x}_1(t)$ and $\dot{x}_2(t)$. The coupled differential equations describing the dynamics of this micromechanical model, including the active feedback loop, may be written in terms of $x_1(t)$ and $x_2(t)$, as

$$\ddot{x}_1 = \frac{1}{M_1}\{p(t) + g\gamma[C_4(\dot{x}_2(t) - \dot{x}_1(t)) + K_4(x_2(t) - x_1(t))] - \dot{x}_1(t)(C_1 + C_3) - x_1(t)(K_1 + K_3) + \dot{x}_2(t)C_3 + x_2(t)K_3\}, \quad (30)$$

$$\ddot{x}_2 = \frac{1}{M_2}\{-\dot{x}_2(t)(C_2 + C_3) - x_2(t)(K_2 + K_3) + \dot{x}_1(t)C_3 + x_1(t)K_3\}. \quad (31)$$

These equations may be written in the form

$$\dot{\mathbf{x}}_n(t) = \mathbf{A}_n\mathbf{x}_n(t) + \mathbf{B}_n p_n(t), \quad (32)$$

by defining the state variable vector for this nth micromechanical element to be $\mathbf{x}_n$, which is equal to $[\dot{x}_1\ x_1\ \dot{x}_2\ x_2]_n^T$. The pressure difference across the cochlear partition, $p(t)$ in Eq. (29), is $p_n(t)$ in this case and $\mathbf{A}_n$ and $\mathbf{B}_n$ are defined as

$$\mathbf{A}_n = \begin{bmatrix} -\frac{(C_1 + C_3 - g\gamma C_4)}{M_1} & -\frac{(K_1 + K_3 - g\gamma K_4)}{M_1} & \frac{C_3 - \gamma C_4}{M_1} & \frac{K_3 - \gamma K_4}{M_1} \\ 1 & 0 & 0 & 0 \\ \frac{C_3}{M_2} & \frac{K_3}{M_2} & -\frac{(C_2 + C_3)}{M_2} & -\frac{(K_2 + K_3)}{M_2} \\ 0 & 0 & 1 & 0 \end{bmatrix}_n, \quad (33)$$

$$\mathbf{B}_n = \left[\frac{1}{M_1}\ 0\ 0\ 0\right]_n^T. \quad (34)$$

The differential equation used by Neely and Kim (1986) to describe the middle ear dynamics can be written

$$p_1(t) = M_S\ddot{w}_{SR}(t) + C_S\dot{w}_{SR}(t) + K_S w_{SR}(t), \quad (35)$$

where M_s, C_s, and K_s are the lumped mass, damper, and stiffness. Defining $\dot{w}_{SR}(t)$ and $w_{SR}(t)$ to be the two state variables for this model, it can be written in the form

$$\dot{\mathbf{x}}_1(t) = \mathbf{A}_1\mathbf{x}_1(t) + \mathbf{B}_1 p_1(t), \quad (36)$$

where

$$\mathbf{x}_1(t) = [\dot{w}_S(t)\ w_S(t)]^T, \quad (37)$$

$$\mathbf{A}_1 = \begin{bmatrix} -\frac{C_s}{M_s} & -\frac{K_s}{M_s} \\ 1 & 0 \end{bmatrix}, \quad (38)$$

$$\mathbf{B}_1 = \left[\frac{1}{M_S}\ 0\right]^T. \quad (39)$$

The pressure release boundary conditions at the helicotrema can also be represented in the state space form of Eq. (32), where

$$x_N = \dot{w}_H, \quad (40)$$

$$A_N = 0, \quad (41)$$

$$B_N = 0. \quad (42)$$

Provided the system is stable, the frequency response of a system described by the state space equations

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t), \quad (43)$$

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t), \quad (44)$$

can be derived by assuming each element of $\mathbf{x}(t)$ is proportional to $e^{j\omega t}$, so that $\mathbf{x}(t)$ can be written as $\mathbf{x}(j\omega)$ and $\dot{\mathbf{x}}(t)$ as $j\omega\mathbf{x}(j\omega)$. The overall matrix of frequency responses can thus be expressed as

$$\mathbf{y}(j\omega) = [\mathbf{D} + \mathbf{C}(j\omega\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}]\mathbf{u}(j\omega). \quad (45)$$

Equation (44) is the general equation for any output quantity, $\mathbf{y}(t)$, as a function of the states, $\mathbf{x}(t)$ and input $\mathbf{u}(t)$. For the case of interest here, the outputs are the BM displacements along the length of the cochlea, which are already contained, as x_{1n}, in the state vector of the micromechanical model, Eq. (32). Thus in this case, the $\mathbf{C}$ matrix just selects the appro-

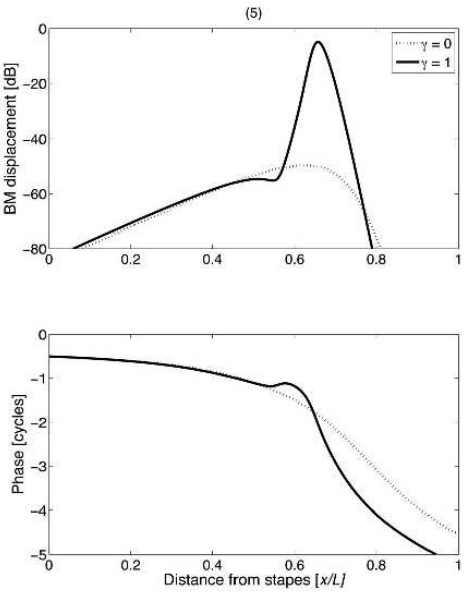


FIG. 5. Distribution of BM displacement along the cochlea calculated from the state space model for the parameters of Neely and Kim (1986) at 1.6 kHz with feedback gain $\gamma=0$ (dashed) and $\gamma=1$ (solid).

priate state variables and the **D** matrix is zero.

Equation (45) has been used with the formulation above and the parameters given in Table I to calculate the BM displacement at various frequencies as a function of position along the modeled cochlea. An example response is shown in Fig. 5, which is equivalent to Fig. 4 in Neely and Kim (1986).

The enhancement of the modeled active response can be varied in the Neely and Kim model by changing the value of the feedback gain, γ. The maximum overall enhancement in the frequency response at each position along the cochlea is plotted in Fig. 6 for the value $\gamma=1$, originally used by Neely and Kim. It can be seen that up to 90 dB enhancement is produced with this level of feedback gain. It has been argued, by Patazzi *et al.* (1989), for example, that this degree of enhancement overestimates that measured in the cochlea. Such a high degree of enhancement also indicates that the

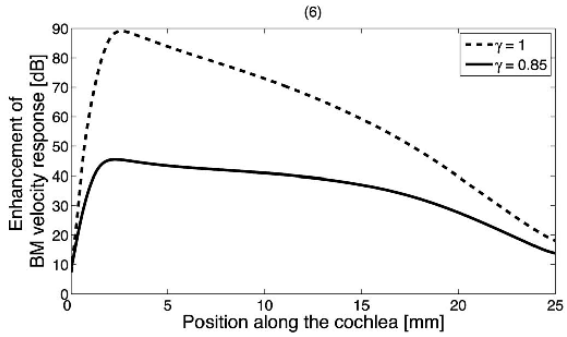


FIG. 6. Enhancement of the peak response in the BM velocity due to the cochlear amplifier for feedback gains of $\gamma=1$ and $\gamma=0.85$.

model is very close to instability (Franklin *et al.*, 2005). The stability of such a model might thus be considered unrealistically sensitive to perturbations in the model's parameters to be representative of the cochlea.

For these reasons, a value of feedback gain of $\gamma=0.85$ has been chosen to be the benchmark for the perturbation studies in this paper. The enhancement produced with this value of feedback gain is also shown in Fig. 6, which indicates that the maximum enhancement due to the cochlear amplifier is about 45 dB near the base and 20 dB near the apex, which is similar to the measured difference in response between live and dead cochlea (as reviewed, for example, by Robles and Ruggero, 2001).

Analyzing the stability of the coupled state space formulation of the cochlea is now readily accomplished by using the state space system to calculate the eigenvalues of the system matrix, **A**, as formulated in Eq. (43). Figure 7 shows these eigenvalues for the coupled system with feedback gains of $\gamma=0$, 0.85, 1.00, and 1.03. In this case, the number of elements in the discretized cochlear model was set to 500, as in the original Neely and Kim (1986) model, which includes the elements that account for the boundary conditions at each end. There are four poles per cochlear element, and three additional poles for the model of the boundary conditions, which makes a total of 1995 poles. Although individual poles can be distinguished on the bottom right-hand side of Fig. 7(b), the majority of the pole positions are so close together that they appear as a continuous line. Note that the distribution of poles in the lower part of Fig. 7(a) extends linearly below the range plotted, which was chosen to retain the detail in the other plots.

The coupled system is stable (since there are no poles with positive real parts) for both the passive model, with $\gamma=0$, and the active model with $\gamma=0.85$ or $\gamma=1$. As the model becomes active, the distribution of the poles nearest the frequency axis, which are the most lightly damped poles, moves further towards the frequency axis, and are so even more lightly damped. They are also stretched out in frequency, reflecting the increase in peak frequency for the active cochlea. When the gain in all the micromechanical elements is increased to 1.03, however, a number of poles appear with positive real parts, indicating that the system is unstable. As well as providing a convenient tool for establishing stability with uniform parameter distributions, the state space model also allows the stability of the cochlear model with nonuniform parameter distributions to be readily calculated, as considered in the following section.

IV. EFFECT OF SPATIAL INHOMOGENEITIES

It has been previously reported that variations in the physical parameters describing the basilar membrane along its length can result in cochlear instability. Although such spatial inhomogeneities could be present in any of the physical parameters in the micromechanical model, we use variations in feedback gain to illustrate some of the effects such inhomogeneities may have on the stability of the coupled system.

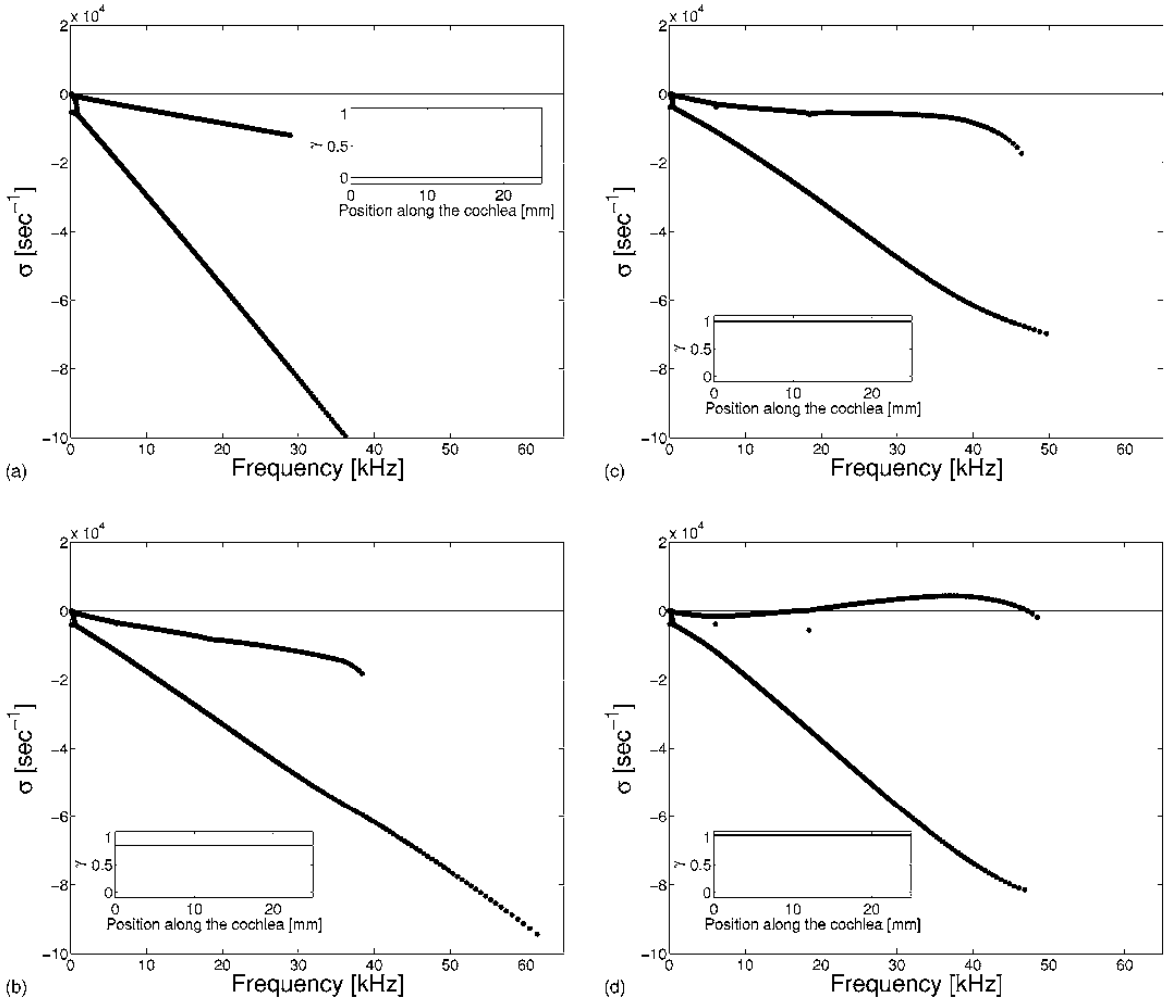


FIG. 7. Distribution of poles in the coupled cochlear model for various values of feedback gain in the micromechanical model, γ. Any pole with a positive real part denotes an unstable system. The cochlear model is clearly stable when it is passive, γ=0 (a), and also for feedback gains of 0.85 and 1 (b) and (c). The cochlear model is unstable if the feedback gain is increased to 1.03 (d).

A. Step changes in gain

Sharp discontinuities in the spatial distribution of gain can result in instability, even if the magnitudes of the changes are small. This is most evident when a step change in gain is introduced. Figure 8(a) shows the pole positions of a cochlea model with a 0.1% step reduction in gain at x =5 mm. The variation of feedback gain (γ) with position along the cochlear is also shown for reference.

In this case, a single pair of poles has a positive real part, indicating that the system is unstable. The frequency of the poles with the positive real part is close to the characteristic frequency of the position on the BM where the discontinuity is located. A 0.1% step increase of gain at 5 mm yields a similar result. The magnitude of the step change required to destabilize the cochlear model varies when discontinuities are introduced at different positions along the cochlea. The system's stability is most sensitive to discontinuities at the basal end, as illustrated above, and discontinuities of up to about 10% can be tolerated near the apical end, before the system becomes unstable.

It is generally accepted that cochlear instability, stabilized by nonlinearities into a limit cycle, is the source of spontaneous otoacoustic emissions (Bialek and Wit, 1984; Talmadge and Tubis, 1993; Nutall *et al.*, 2004). One theory regarding SOAEs suggests that they are not produced locally, but rather are the result of global wave scattering and active reinforcement of reflections by the active process, likening the cochlea to a laser oscillator (Talmadge and Tubis, 1993; Shera, 2003). A laser oscillator includes a resonant cavity which encloses a gain medium. The cochlear equivalent of a laser's resonant cavity exists between the region of a spatial discontinuity and the middle ear boundary; within that area, a minimum amount of energy must be reflected at each boundary in order to sustain a limit cycle. Here the outer hair

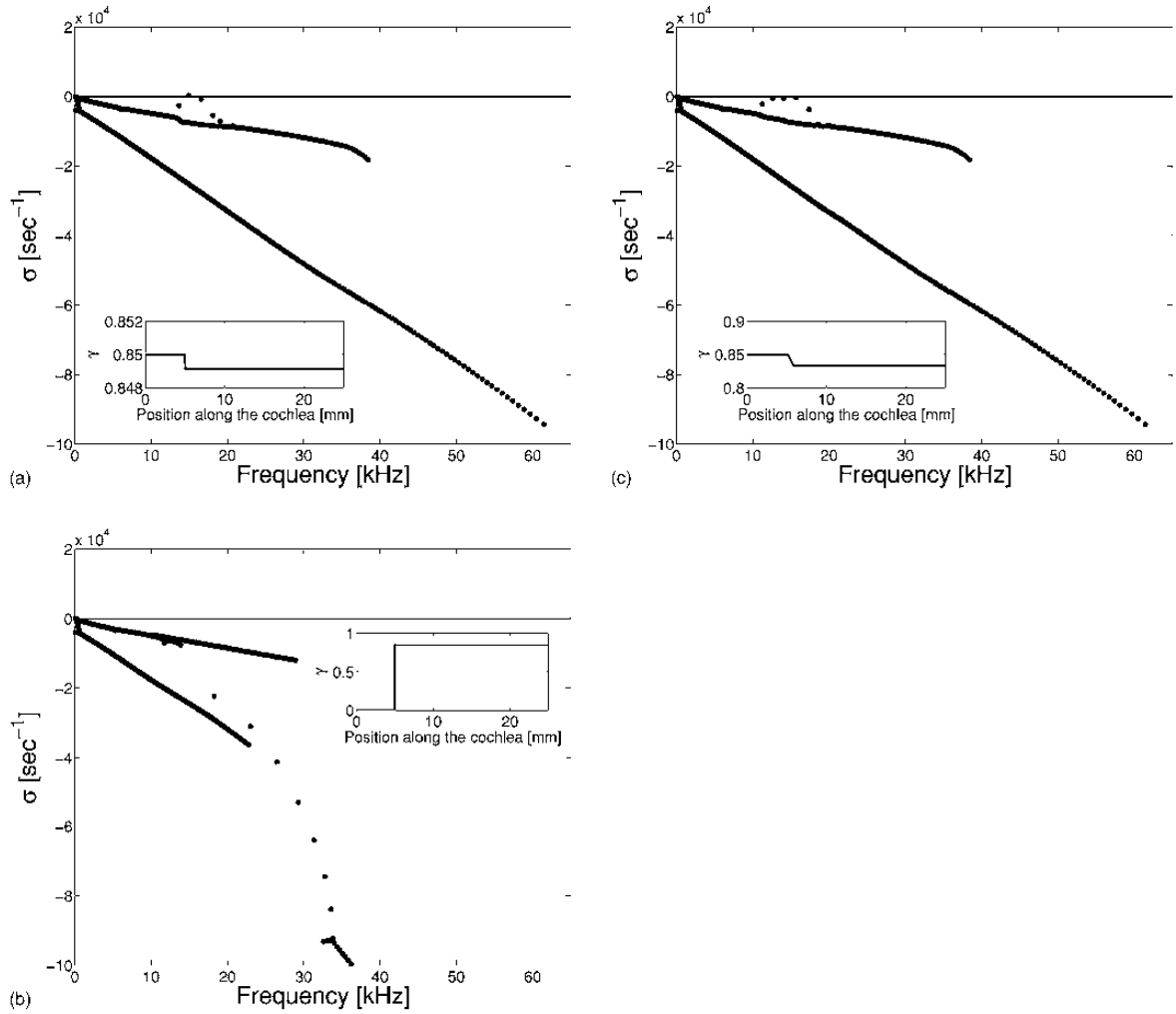


FIG. 8. Poles of the coupled cochlea for a gain distribution of feedback gains with: (a) a 0.1% step decrease at $x=5$ mm; system is unstable; (b) a step gain distribution of magnitude 0.85 at $x=5$ mm; system is stable; and (c) a smoothly varying gain distribution (quarter sin-squared wave dip) of magnitude 2% change at $x=5$ mm; system is stable. Note that the sketches of feedback gain with position along the cochlea have differing scales.

cells (OHCs) of the cochlea amplify waves traveling in each direction. While SOAEs are normally thought to be generated by random spatial perturbations in the dynamics of the cochlea (Zweig and Shera, 1995; Shera, 2003), pathologically abrupt discontinuities are also recognized as generating "atypical SOAEs" (Clark *et al.*, 1984; Lonsbury-Martin and Martin, 2001). Strictly speaking, it is oscillations of the BM motion that are being predicted here rather than the SOAEs, and modeling the detailed relationship between these oscillations and pressure in the external ear is the subject of current research.

Though small discontinuities in gain can drive the coupled cochlea unstable, one extreme spatial discontinuity in gain that does not generate instability is shown in Fig. 8(b), in which the cochlear amplifier gain is set to zero for all positions basal of $x=5$ mm. Note that the poles in the lower part of the graph continue below the plotted range, as in Fig. 7(a), and are very similar to those of the passive system in this frequency range, as expected. In light of the aforementioned theory of SOAE production, the system shown in Fig. 8(b) would have a zero roundtrip gain for any reflected waves between the gain discontinuity and the middle ear of the cochlea. As a result, this would be predicted not to produce enough energy to overcome the losses in that region, and consequently, no instability or SOAE would be produced; this is what is observed in the model here. Contrary to this observation, Fukazawa and Tanaka (1996) found that a similar gain distribution did result in system instability in their time domain model. It may be the feed-forward implementation of their cochlear model that produces this result.

B. Smooth variation in gain

Following the line of reasoning outlined above, it is possible that smoothly varying distributions of gain, associated with the compression of the cochlear response by a low pass

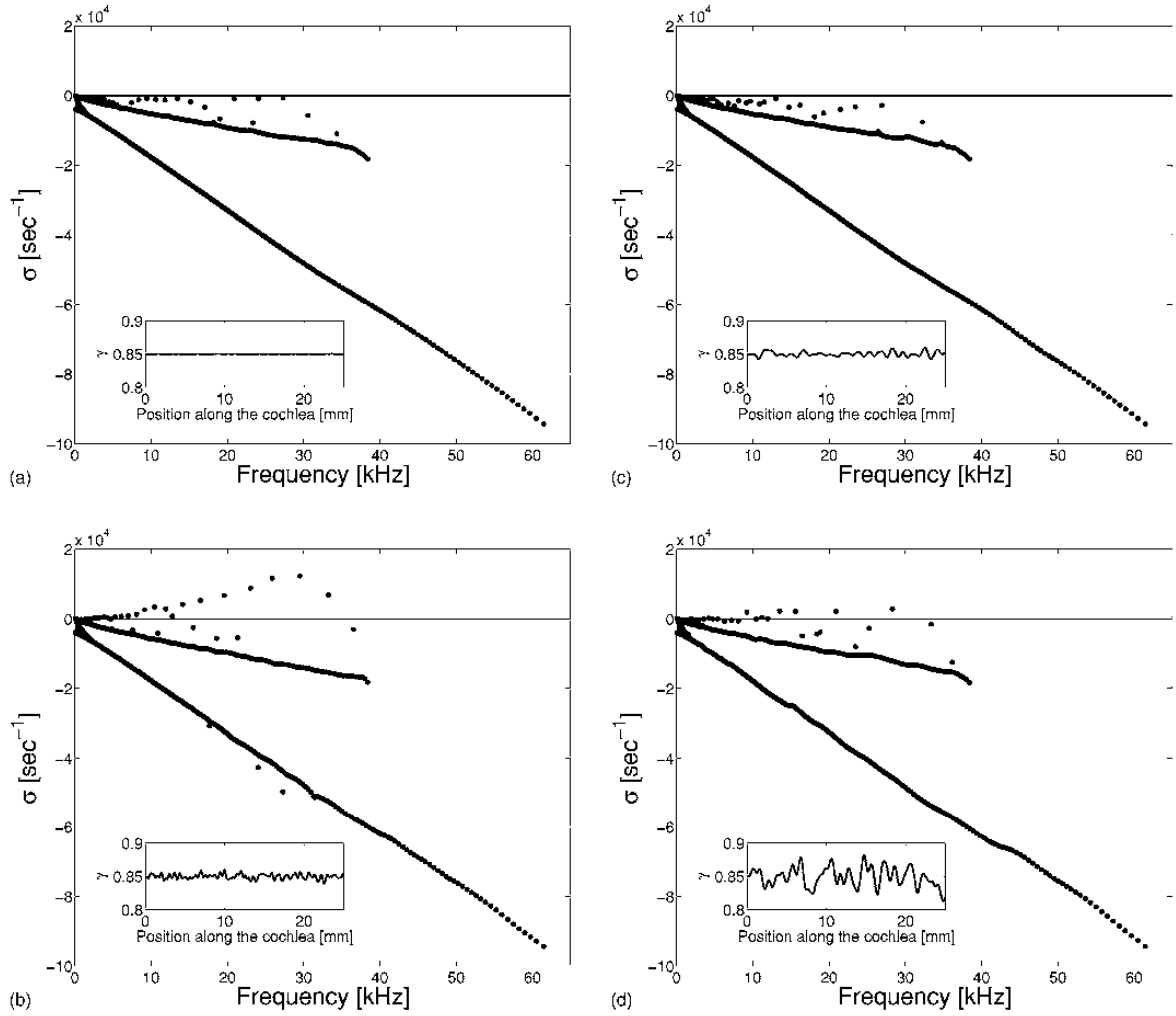


FIG. 9. Poles of the coupled cochlea for a random gain distribution of feedback gains: (a) with small magnitude (0.1%) and rough spatial variations; system is stable; (b) intermediate magnitude (~2%) and rough spatial variations; system is unstable; (c) intermediate magnitude (~2%) and smooth spatial variations; system is stable; and (d) large magnitude variations (~8%) and smooth spatial variations; system is unstable.

signal for example, might not reflect enough energy to sustain limit cycle oscillations. Figure 8(c) shows the pole positions for a 2% dip in gain over a 1 mm region of the cochlea centered on $x=5$ mm. The spatial distribution in the gain used in Fig. 8(c) is proportional to a quarter-$\sin^2(x)$ wave, and proves smooth enough to maintain stability. In contrast, a quarter-$\sin(x)$ wave form dip of the same magnitude in the same position, whose edges are more abrupt than the quarter-$\sin^2(x)$ dip, does not maintain stability. With the smooth quarter-$\sin^2(x)$ variation, even step changes of 100% in gain at the apical end of the cochlea do not destabilize the system.

V. SPATIALLY RANDOM VARIATION IN GAIN

Multiple spontaneous oscillations were observed in a nonlinear hardware model of a section of the cochlea by Zwicker (1986). Previous time domain simulations of active cochlear models have also produced internal oscillations that are believed to be representative of SOAEs (Fukazawa and Tanaka, 1996; van Hengel *et al.*, 1996; Talmadge and Tubis 1998). The active process has either been modeled using a feed-forward active force that is due to the incoming velocity from the base (Fukazawa and Tanaka, 1996) or a negative linear damper (van Hengel *et al.*, 1996; Talmadge and Tubis 1998).

In this section, linear systems with the Neely and Kim micromechanical model having randomly varying gain values are tested for stability. An entirely random wave form with a Gaussian distribution was band-pass filtered at different spatial frequencies in order to give smoothly or roughly spatially varying gain distribution. The filter applied is a fifth-order Butterworth filter, chosen for its characteristically flat passband (Lineton, 2001). The lower boundary of the filter is set at a spatial frequency given by the reciprocal of

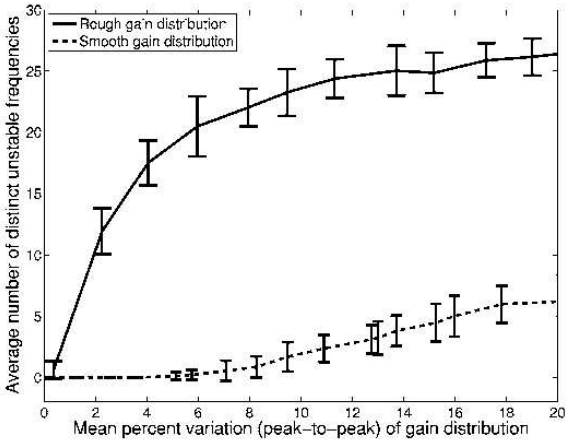


FIG. 10. The number of unstable poles in coupled cochlea models having different amplitudes of cochlear amplifier gain variation with either smooth or rough distributions along its length. The bars denote the range of 40 simulations with different random gain distributions and the circles denote the average value.

the length of the cochlea. The upper spatial frequency boundary of the filter is set within a range of 5–10% of the spatial sample rate, which corresponds to wavelengths on the order of several millimeters down to a fraction of 1 mm. Two upper spatial frequencies were used here: equivalent to wavelengths of 0.85 mm, which corresponds to a "rough" gain function and 1.7 mm, which corresponds to a "smooth" spatially varying gain function.

Figure 9 shows examples of randomly varying gain functions and their corresponding pole distributions. It can be seen that an increased magnitude of spatially varying gain fluctuation will cause instability, although lower percentage variations are required to cause the system to become unstable for rough spatially varying gains.

Different random gain variations will produce different distributions of poles, but the average number of unstable poles is dependent on both the smoothness of the gain distribution and its magnitude. This behavior was investigated in a series of simulations whose results are summarized in Fig. 10. The range of the number of unstable poles observed in 40 simulations is shown as a bar for different peak-to-peak percentage variations in gain. The average number of unstable poles is also shown, and these points are connected to illustrate the way in which the average number of unstable poles increases with percent variation in gain for both the rough and the smooth gain distributions. The system with the rough gain distribution becomes unstable even for 0.5% variations, whereas the system with the smooth gain distribution is stable for variations of up to about 5%. The average spacings between the poles when they are close to instability also cluster around a fixed value, as predicted by the theory of Zweig and Shera (1995), although the details of this behavior are still under investigation and will be the subject of a future publication.

VI. NONLINEAR TIME DOMAIN SIMULATIONS

Although an eigenvalue analysis of the state matrix can reveal whether a linear model is unstable, the exponentially

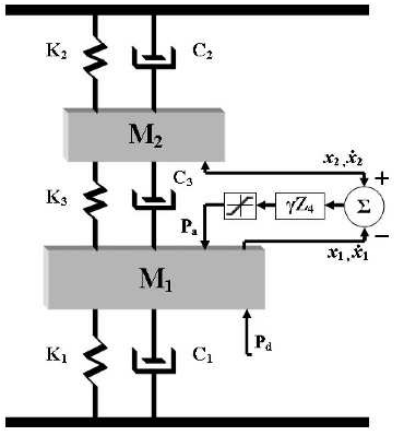


FIG. 11. Block diagram of the micromechanical model of Neely and Kim (1986) with a saturating nonlinearity in the active force.

diverging response of such a linear unstable system in a real cochlea would be limited by nonlinearities. In order to explore the complex behavior due to such nonlinearities, it is necessary to analyze the system in the time domain. Figure 11 shows the Neely and Kim micromechanical model with a saturation function operating on the active force, which might represent the nonlinearity of the outer hair cell response. This saturation is modeled here as a tanh function, which is linear for low-level excitation. It was noted in Sec. I that if the feedback gain, γ, was set to 1.1 in such an isolated model, the linear system would be unstable. The nonlinearity introduced in Fig. 11 will limit the amplitude of such an instability, as illustrated in Fig. 12, which shows the results of a time domain simulation of such an isolated nonlinear micromechanical element that is linearly unstable. The nonlinearity limits the amplitude of the oscillation, as predicted, so that in this case it settles down into a periodic limit cycle.

The state space formulation provides a convenient tool for such time domain analysis in the coupled cochlea model, using, for example, ordinary differential equation (ODE) solvers such as ode45 in MATLAB, which uses a Runge–Kutta

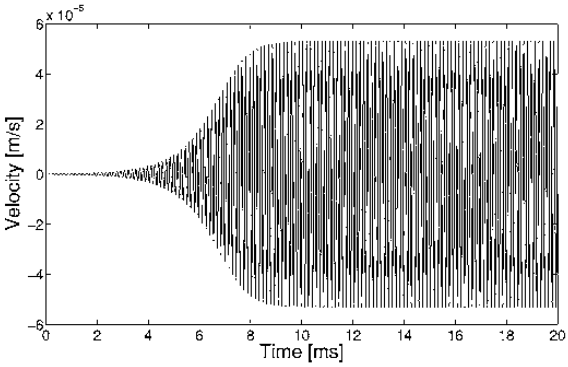


FIG. 12. The results of a time domain simulation of a single isolated nonlinear micromechanical element with γ=1.1, which is linearly unstable. The amplitude of the oscillation initially diverges exponentially, but is then limited by the saturation function so that it settles into a periodic limit cycle.

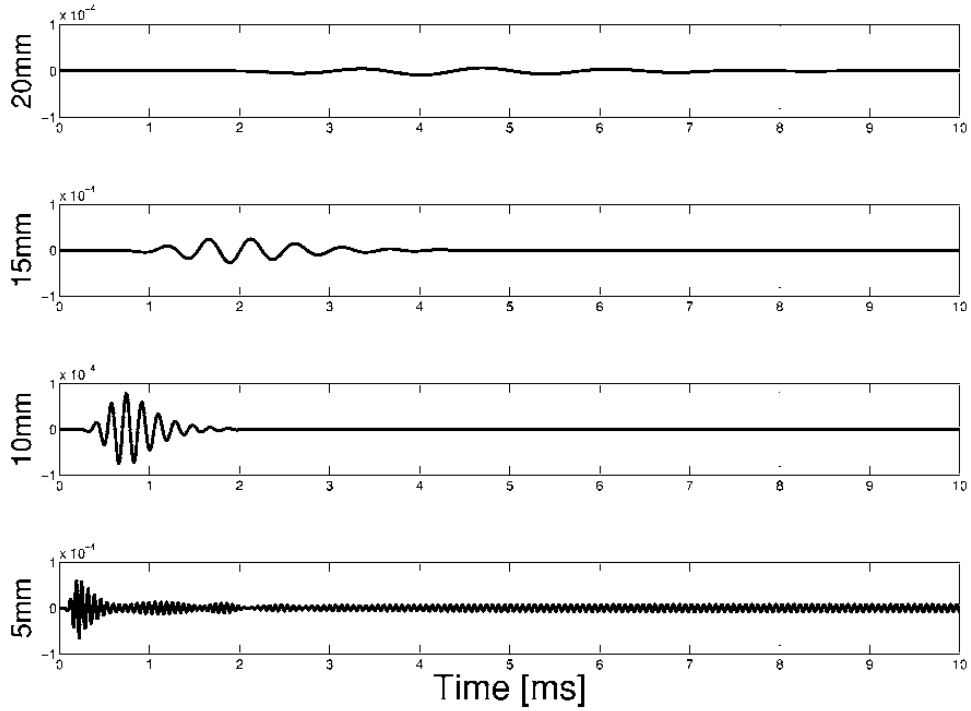


FIG. 13. BM velocity responses at different positions along the coupled cochlea as a result of a time domain simulation with saturating cochlear amplifiers and a distribution of feedback gains corresponding to Fig. 8(a).

algorithm with variable internal step size. This program has been used to simulate Eq. (26) in the time domain with an output sample rate of 100 kHz, which is necessary to avoid aliasing at the higher characteristic frequencies. The error tolerance in the ODE solver was set to 1e-10. Extremely similar results were obtained when the error tolerance was set to 1e-11, although the computation time was considerably greater in this case. The time domain simulations were initially validated against the frequency domain results by calculating the "steady-state" response at all positions along the cochlea to a 6 kHz tone after 100 ms of simulation. The results were almost identical to the frequency domain results, except where the response had fallen by about 200 dB at the apical end, in which case the finite window of the time domain simulations caused some spurious results.

Time domain simulations have also been performed with a nonlinear active model having a 0.1% step in the spatial gain distribution at $x=5$ mm, as used in Fig. 8(a), which is just sufficient to cause a single pair of poles to have positive real parts and is thus linearly unstable. The nonlinearity is incorporated into the simulation by replacing γ in the $\mathbf{A}_n$ matrix for each active micromechanical element, Eq. (33), with the instantaneous ratio of the input to output of the tanh function for that element. Figure 13 shows the evolution of the BM response at a number of positions along the cochlea for this gain distribution along the cochlea. In order to reduce the simulation time before the nonlinear model reaches a steady state, the cochlea is initially excited by a short pulse at the basal end. This causes a transient that dies away within 2 ms for positions less than about 10 mm along the cochlea, but takes longer to die away at more apical positions because of dispersion. This is more clearly seen in the grey scale plot shown in Fig. 14, which indicates the amplitude of the cochlear pressure as a function of time at all modeled positions along the cochlea. The original transient due to the impulsive excitation corresponds to an arc starting at the origin and

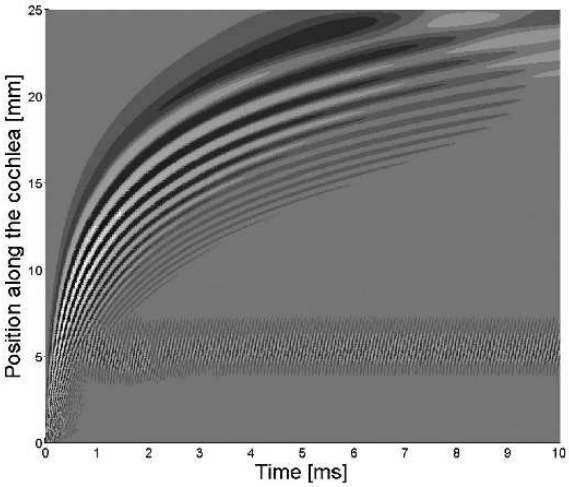


FIG. 14. Grey scale representation of the results of the time domain simulation described in Fig. 13 at all positions along the cochlea. The grey scale corresponds to the amplitude of the pressure (positive white, negative black) at each position at each time. An animation of this response is available online (Ku, 2007).

spreading out to 15 mm after about 3 ms. The strongest response after this transient has died away, however, is a single oscillation centered at $x=5$ mm along the cochlea. The frequency of this limit cycle is very similar to that of the imaginary part of the unstable pole, which is itself very similar to the characteristic frequency at the position of the step in gain.

VII. CONCLUSIONS

The main contribution of this paper is the formulation of a general state space model of the coupled cochlea. The prevalence of analytical and numerical tools for dynamic systems cast in this time domain form then allows several aspects of the behavior of cochlear models to be explored in more detail than is possible with an entirely frequency domain model.

Particular emphasis is placed on the assessment of stability for a linear model. The results of frequency domain models cannot be trusted unless the system can be demonstrated to be stable, and the pole positions derived from the state space model provide a convenient way of establishing whether such a coupled linear model is stable.

The stability of the cochlear model is illustrated using the model of Neely and Kim (1986). In particular, the effect of different distributions of micromechanical feedback gains along the cochlea is considered. It is shown that abrupt changes in feedback gain can cause instability even when the magnitude of the change is less than 0.1%. Smoother changes in gain can be much larger in magnitude before the stability of the model is compromised. The effect of various spatially random variations of feedback gain on stability is also investigated. It is found that there is a significant difference in the magnitude of such inhomogeneities that cause instability, depending on the smallest spatial wavelength of the feedback gain distribution.

Finally, we illustrate the use of the state space formulation to simulate the coupled system. Nonlinearities in the micromechanical feedback loops are included in order to begin to account for the finite active contribution of the OHCs. The limit cycle oscillations that are generated when the linearly unstable response is limited by the saturating nonlinearity are thought to be the source of SOAEs. The state space formulation may also be a convenient tool to study other forms of otoacoustic emission.

ACKNOWLEDGMENTS

We thank Professor Stephen Neely for his help in defining the parameters in his 1986 model. The work of Emery M. Ku was supported by a Fulbright Postgraduate Award.

Allen, J. B. (**1977**). "Cochlear micromechanics—a mechanism for transforming mechanical to neural tuning within the cochlea," J. Acoust. Soc. Am. **62**(4), 930–939.

Bialek, W., and Wit, H. (**1984**). "Quantum limits to oscillator stability: Theory and experiments on acoustic emissions from the human ear," Phys. Lett. **104A**(3), 173–178.

Clark, W. W., Kim, D. O., Zurek, P. M., and Bohne, B. A. (**1984**). "Spontaneous otoacoustic emissions in chinchilla ear canals: Correlation with histopathology and suppression by external tones," Hear. Res. **16**, 299–314.

de Boer, E. (**1996**). "Mechanics of the cochlea: modelling effects," *The Cochlea*, edited by P. Dallas, A. N. Popper, and R. R. Ray, Chap. 5 (Springer-Verlag, New York, NY).

Franklin, G. F., Powell, J. D., and Emani-Naeini, A. (**2005**). *Feedback Control of Dynamics Systems*, 5th ed. (Addison Wesley, Reading, MA).

Fukazawa, T., and Tanaka, Y. (**1996**). "Spontaneous otoacoustic emissions in an active feed-forward model of the cochlea," Hear. Res. **95**, 135–143.

Furuta, K., Sano, A., and Atherton, D. (**1988**). *State Variable Methods in Automatic Control* (Wiley, New York).

Kanis, L. J., and de Boer, E. (**1993**). "Self-suppression in a locally active non-linear model of the cochlea: A quasilinear approach," J. Acoust. Soc. Am. **94**, 3199–3206.

Koshigoe, S., and Tubis, A. (**1983**). "Frequency-domain investigations of cochlear stability in the presence of active animals," J. Acoust. Soc. Am. **73**(4), 1244–1248.

Ku, E. M., ed. (**2007**). *Animations of time domain simulations*, Institute of Sound and Vibration Research, University of Southampton. Last viewed 8/14/2007. ⟨http://www.isvr.soton.ac.uk/SPCG/Cochlear_Mechanics/JASA2007/index.html⟩

Lindgren, A. G., and Li, W. (**2003**). "Analysis and simulation of a classic model of cochlea mechanics via a state-space realization." Last viewed 8/14/2007. ⟨http://www.ele.uri.edu/~liw/bm.pdf⟩

Lineton, B. (**2001**). "Testing a model of the stimulus frequency otoacoustic emissions in humans," PhD thesis, University of Southampton, Southampton, UK.

Lonsbury-Martin, B. L., and Martin, G. K. (**2001**). "Otoacoustic emissions," *Physiology of the Ear*, edited by A. F. Jahn and J. Santos-Sacchi (Singular, San Diego), 443–480.

Monderer, B., and Lazar, A. A. (**1988**). "Speech signal detection at the output of a cochlear model," Proceedings of International Conference on Acoustics, Speech, and Signal Processing, ICASSP-88, 11–14 April, 1988, New York, NY.

Neely, S. T. (**1981**). "Finite difference solution of a two-dimensional mathematical model of the cochlea," J. Acoust. Soc. Am. **69**, 1386–1393.

Neely, S. T., and Kim, D. O. (**1986**). "A model for active elements in cochlear biomechanics," J. Acoust. Soc. Am. **79**(5), 1472–1480.

Nuttall, A. L., Grosh, K., Zheng, J., de Boer, E., Zou, Y., and Ren, T. (**2004**). "Spontaneous basilar membrane oscillation and otoacoustic emission at 15 kHz in a guinea pig," J. Assoc. Res. Otolaryngol. **5**, 337–348.

Patuzzi, R. B., Yates, G. K., and Johnstone, B. M. (**1989**). "Outer hair cell receptor current and sensorineural hearing loss," Hear. Res. **42**(1), 47–72.

Robles, L., and Ruggero, M. A. (**2001**). "Mechanics of the mammalian cochlea," Physiol. Rev. **81**(3), 1305–1352.

Shera, C. A. (**2003**). "Mammalian spontaneous otoacoustic emissions are amplitude-stabilized cochlear standing waves," J. Acoust. Soc. Am. **114**(1), 244–262.

Shera, C. A., and Zweig, G. (**1993**). "Order from chaos: Resolving the paradox of periodicity in evoked otoacoustic emissions," *Biophysics of Hair-Cell Sensory Systems*, edited by H. Duifhuis, J. W. Horst, P. van Dijk, and S. M. van Netten (World Scientific, Singapore) 54–63.

Skogestad, S., and Postlethwaite, I. (**1996**). *Multivariable Feedback Control* (Wiley, New York).

Stibler, B. Z., Lewis, E. R., and Henry, K. R. (**1998**). "A state space model of gerbil cochlea," Proceedings of the 6th Annual Conference on Computational Neuroscience: Trends in research, 26–30 July, 1998, Big Sky, MT, 107–112.

Talmadge, C. L., and Tubis, A. (**1993**). "On modelling the connection between spontaneous and evoked otoacoustic emissions," *Biophysics of Hair Cell Sensory Systems*, edited by H. Duifhuis, J. W. Horst, P. van Dijk, and S. M. van Netten (World Scientific, Singapore), pp. 25–32.

Talmadge, C. L., and Tubis, A. (**1998**). "Modeling otoacoustic emission and hearing threshold fine structures," J. Acoust. Soc. Am. **104**(3), 1517–1543.

van Hengel, P. W. J., Dulfuis, 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**(6), 3566–3571.

Zweig, G. (**1991**). "Finding the impedance of the organ of corti," J. Acoust. Soc. Am. **89**, 1229–1254.

Zweig, G., and Shera, C. A. (**1995**). "The origin of periodicity in the spectrum of evoked otoacoustic emissions," J. Acoust. Soc. Am. **98**(4), 2018–2047.

Zwicker, E. (**1986**). "A hardware cochlear model with active feedback," J. Acoust. Soc. Am. **80**(1), 146–153.

Zwislocki, J. (**1950**). "Theory of the acoustical action of the cochlea," J. Acoust. Soc. Am. **22**, 778–784.

C.2 Ku et al. *2008*

The text, figures and simulations presented in this journal paper were developed by E.M. Ku. Guidance and editorial assistance were provided by S.J. Elliott and B. Lineton.

Appendix C

Statistics of instabilities in a state space model of the human cochlea

Emery M. Ku,[a] Stephen J. Elliott, and Ben Lineton
Institute of Sound and Vibration Research, University of Southampton, Southampton, United Kingdom SO17 1BJ

(Received 16 January 2008; revised 8 May 2008; accepted 8 May 2008)

A state space model of the human cochlea is used to test Zweig and Shera's [(1995) "The origin of periodicity in the spectrum of evoked otoacoustic emissions," J. Acoust. Soc. Am. **98**(4), 2018–2047] multiple-reflection theory of spontaneous otoacoustic emission (SOAE) generation. The state space formulation is especially well suited to this task as the unstable frequencies of an active model can be rapidly and unambiguously determined. The cochlear model includes a human middle ear boundary and matches human enhancement, tuning, and traveling wave characteristics. Linear instabilities can arise across a wide bandwidth of frequencies in the model when the smooth spatial variation of basilar membrane impedance is perturbed, though it is believed that only unstable frequencies near the middle ear's range of greatest transmissibility are detected as SOAEs in the ear canal. The salient features of Zweig and Shera's theory are observed in this active model given several classes of perturbations in the distribution of feedback gain along the cochlea. Spatially random gain variations are used to approximate what may exist in human cochleae. The statistics of the unstable frequencies for random, spatially dense variations in gain are presented; the average spacings of adjacent unstable frequencies agree with the preferred minimum distance observed in human SOAE data. © *2008 Acoustical Society of America.* [DOI: 10.1121/1.2939133]

PACS number(s): 43.64.Kc, 43.64.Jb, 43.40.Vn, 43.64.Bt [BLM] Pages: 1068–1079

I. INTRODUCTION

The existence of a cochlear amplifier (CA) was first postulated by Gold (1948), who argued an electromechanical action is necessary to counteract the heavy viscous damping in the fluid-filled cochlea. The discovery of spontaneous otoacoustic emissions (SOAEs) by Kemp (1979) has long served as indirect evidence supporting the presence of a CA. It is now widely accepted that the outer hair cells situated in the organ of Corti actively enhance the motion of the basilar membrane (BM) (e.g., Diependaal *et al.*, 1987), which gives rise to a mammal's sharply tuned sense of hearing. However, the precise mechanism underlying the generation of SOAEs is still in debate.

SOAEs are believed to be a feature of a normally functioning CA, as they are commonly detected in an estimated range of 33%–70% of all normally hearing ears (Talmadge *et al.*, 1993). Where SOAEs are detected, stimulus frequency-, distortion product- and transient evoked-otoacoustic emissions (SFOAEs, DPOAEs, and TEOAEs) are often present. There is evidence to suggest that all forms of OAEs are related and directly tied to the sensitivity of hearing (Zwicker and Schloth, 1984; McFadden and Mishra, 1993; Talmadge and Tubis, 1993; Shera and Guinan, 1999). Two primary classes of cochlea-based theories regarding the production of SOAEs are discussed below: a local-oscillator model and a distributed backscattering concept.

Gold (1948) first formed the basis of a local-oscillator model of SOAE generation when he proposed that a perturbation may "bring an [active] element into the region of self-oscillation, when it is normally so close to [instability]." Evidence in the literature suggests that SOAEs are associated with BM oscillations. For example, Nuttall *et al.* (2004) measured a SOAE that had a counterpart in spontaneous mechanical vibration of the BM at the same frequency. Further work performed by Martin and Hudspeth (2001) considered how locally unstable elements of the CA may be responsible for SOAEs. However, without careful tuning, a local-oscillator model fails to account for the regular spacings between unstable frequencies observed in mammalian SOAEs.

The strong peak in the distribution of spacings between adjacent SOAE frequencies, termed the preferred minimum distance (PMD), has been demonstrated by various studies (Dallmayr, 1985, 1986; Talmadge *et al.*, 1993; Braun, 1997). A similar value is found in the average frequency spacings between the spectral peaks of SFOAEs and TEOAEs when measured in the ear canal (Zwicker and Schloth, 1984; Shera, 2003). The PMD corresponds to a frequency spacing of approximately 0.4 bark, or a distance of about 0.4 mm along the human cochlea (Dallmayr, 1985, 1986). Most SOAEs occur in the range of 0.5–6 kHz (Probst *et al.*, 1990) and demonstrate the PMD, though Zweig and Shera (1995) and Shera (2003) showed that the average spacings of both SOAEs and the spectral peaks of SFOAEs measured in the ear canal vary somewhat with frequency.

Strube (1989) argued that a periodic variation or "corrugation" in the micromechanical parameters would also give rise to the observed PMD in SFOAE and TEOAE measurements in the ear canal. This was said to arise given distributed backscattering of the traveling wave (TW) similar to the phenomenon of Bragg reflection in a crystal. In this theory, the period of the corrugation must correspond to one-half of

[a]Electronic mail: ek@isvr.soton.ac.uk.

Appendix C

the wavelength of the TW, thus generating constructive interference at particular frequencies. Kemp (1979) also proposed a theory of SOAE generation which assumed a distributed backscattering mechanism; his theory required multiple internal reflections of forward- and backward-traveling waves between the middle ear boundary and an inhomogeneous region of the cochlea.

Since Kemp (1979) first presented the idea, numerous authors have made contributions to the multiple-reflection theory (Zwicker and Peisl, 1990; Zweig, 1991; Shera and Zweig, 1993; Talmadge and Tubis, 1993; Zweig and Shera, 1995; Allen *et al.*, 1995; Talmadge *et al.*, 1998; Shera and Guinan, 1999; Shera, 2003). Shera and Zweig (1993) proposed that a spatially dense and random array of reflection sites exists along the entire cochlea which acts in concert with the middle ear boundary to form standing waves, which Shera (2003) likens to a laser cavity. This concept was fully developed in Zweig and Shera (1995). Though energy is reflected at all frequencies by a perturbation along the cochlea, wavelets scattered from forward-traveling waves that peak in the region of the inhomogeneity dominate the response, since the amplitude is highest there.

For an active standing wave resonance to develop in this multiple-reflection theory, the spatial distribution of inhomogeneities in the given region must contain components at the wavenumber that creates constructive interference with the incoming wave, just as with Bragg scattering (Shera and Zweig, 1993; Zweig and Shera, 1995). Further requirements include an active region between the middle ear boundary and the reflection site to overcome the viscous damping in the cochlea, and a TW frequency that undergoes an integer number of cycles of round-trip phase change between the middle ear and the cochlear reflection site; this naturally gives rise to the PMD in SOAEs measured in the ear canal. However, the existence of a spontaneous oscillation in the cochlea does not guarantee its detection as a SOAE; it must also remain sufficiently powerful to be measurable in the ear canal after transmission through the middle ear.

An alternative theory suggests that irregular middle ear transmission characteristics may be a cause of some OAEs (Nobili *et al.*, 2003). However, the numerical accuracy of these simulation results has been contested elsewhere (Shera *et al.*, 2003), and such irregularities are not often reported. For the purposes of this investigation, a smooth middle ear boundary is implemented and only cochlea-based theories of SOAE generation are discussed.

It should be noted that this paper considers only the linear stability of the cochlear model. In a biological cochlea, the amplitude of an instability would eventually stabilize due to the natural saturation of the feedback force generated by the CA. Furthermore, it is possible that the number of SOAEs predicted by the linear model could change in a nonlinear model due to distortion or suppression, for example.

A. Aims and overview

The goal of this paper is to test whether the predictions formalized by Zweig and Shera's (1995) multiple-reflection theory of SOAE generation are observed in a mathematical model of linear cochlear mechanics. Previous work has relied upon phenomenological methods (Zweig and Shera, 1995; Shera, 2003), or multiple time-domain simulations (Talmadge *et al.*, 1998), to support this theory. In contrast, a state space formulation of the cochlea (Elliott *et al.*, 2007) is used here that is capable of rapidly and unambiguously calculating the unstable frequencies in a given linear model. This method is thus especially well suited to generating the large number of results from individual cochleae necessary to ensure statistically significant data.

Section II presents the revisions necessary to adapt the original model (Neely and Kim, 1986), on which the state space model of Elliott *et al.* (2007) was based, from representing a cat cochlea to representing a human cochlea. For instance, a boundary approximating the human middle ear is now included. The features of the model that are pertinent to the "cochlear laser" theory, such as the wavelength of a TW at its peak as a function of position, are examined. Sample frequency responses and the stability of a base line cochlear model are also briefly described.

In Sec. III, the smoothly varying BM impedance along the cochlea is perturbed with a variety of spatial inhomogeneities in the micromechanical feedback gain in order to introduce reflection sites. The following inhomogeneities are tested: a step change in gain; sinusoidal variations in gain; and band-limited spatially random variations in gain are applied in order to simulate what may exist in human cochleae. A large number of simulations from the last category are performed. The spacings of adjacent unstable frequencies in the randomly perturbed cochlear models are collected and statistically analyzed at the end of this section.

II. MODEL DESCRIPTION

Elliott *et al.* (2007) used a state space formulation to determine the stability of Neely and Kim's (1986) discrete, long wave model of the cat cochlea. The goal of the current work is to be able to compare numerical simulations to human measurements; thus, revisions to the model were necessary to account for the pertinent features of the human cochlea. The changes are described in this section: starting at the middle ear boundary at the oval window, followed by the micromechanical elements of the cochlea, and ending at the helicotrema boundary at the apex. Illustrative simulations and the features of the model pertaining to stability are presented after the revisions.

Shera and Zweig (1990) pointed out the importance of the middle ear boundary as the dominant source of reflections for retrograde TWs in the cochlea. As such, careful attention was given to creating a boundary condition in the revised model that approximates the key features of physiological measurements. The data of Puria (2003) was used as a target when revising Neely and Kim's (1986) mass-spring-damper boundary.

Table I. lists the modified values of the micromechanical elements used in this model, and Fig. 1. shows Z_{out}, the impedance looking out of the cochlea into the middle ear[1] for both the state space model and Puria's (2003) measurements.

Appendix C

TABLE I. Revised parameters of the micromechanical model, as described in Elliott *et al.* (2007), in SI units, where x is the longitudinal distance along the cochlea.

Quantity	Formula (SI)
$k_1(x)$	$4.95\times10^{9}e^{-320(x+0.00375)}$ N m^{-3}
$c_1(x)$	$1+19700\ e^{-179(x+0.00375)}$ N s m^{-3}
m_1	1.35×10^{-2} kg m^{-2}
$k_2(x)$	$3.15\times10^{7}e^{-352(x+0.00375)}$ N m^{-3}
$c_2(x)$	$113\ e^{-176(x+0.00375)}$ N s m^{-3}
m_2	2.3×10^{-3} kg m^{-2}
$k_3(x)$	$4.5\times10^{7}e^{-320(x+0.00375)}$ N m^{-3}
$c_3(x)$	$22.5\ e^{-64(x+0.00375)}$ N s m^{-3}
$k_4(x)$	$2.82\times10^{9}e^{-320(x+0.00375)}$ N m^{-3}
$c_4(x)$	$9650\ e^{-164(x+0.00375)}$ N s m^{-3}
γ	1
H	0.001 m
L	0.035 m
A_s	3.2×10^{-6} m^2
k_{ME}	2.63×10^{8} N m^{-3}
c_{ME}	2.8×10^{4} N s m^{-3}
m_{ME}	2.96×10^{-2} kg m^{-2}
c_H	210 N s m^{-3}
m_H	1.35×10^{-2} kg m^{-2}
N	500

The micromechanical model and the significance of all the quantities are described in Elliott *et al.* (2007). The values of the Neely and Kim's (1986) parameters have been scaled in order to obtain a distribution of characteristic frequencies that matches those of Greenwood (1990) over the range of interest. Whereas Elliott *et al.* (2007) left the boundary at the helicotrema as a pressure release, it is now revised to include a small amount of damping. In order to incorporate the damped boundary into the state space model, it was necessary to make a minor modification to the macromechanical fluid-coupling matrix. The details of the new boundary condition and the revised matrix are explained in the Appendix. This change only affects simulations at low frequencies by reducing the reflectivity of the helicotrema, thus simplifying the interpretation of results.

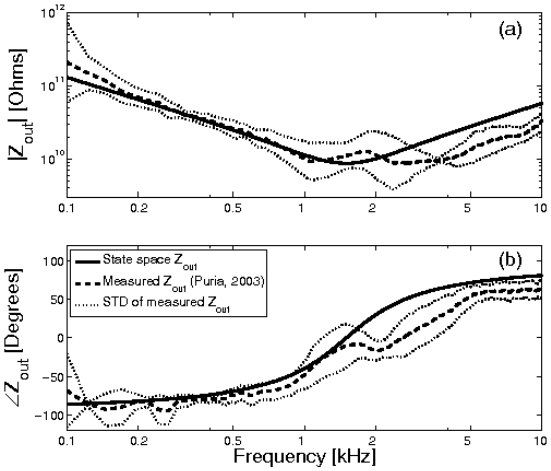


FIG. 1. Magnitude (a) and phase (b) of the impedance of the state space middle ear boundary and measured impedance looking out of the cochlea, Z_{out} (Puria, 2003).

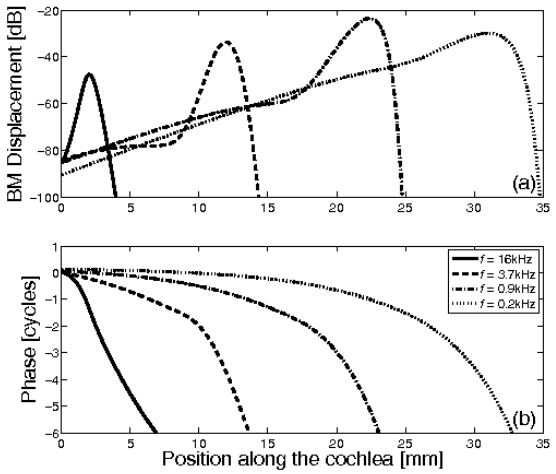


FIG. 2. BM displacement magnitude (a) and phase (b) given the four stimulus tones at f=16, 3.7, 0.9, and 0.2 kHz in the base line model [$\gamma(x)=1$].

The macromechanical formulation of the state space model (Elliott *et al.*, 2007) was based on work by Neely (1981) and Neely and Kim (1986). This uses a finite difference approximation to discretize the spatial derivatives in the wave equation and boundary conditions of the cochlea. The local activity of the cochlear partition segments is related to the fluid mechanics by

$$\mathbf{F}\mathbf{p}(t) - \ddot{\mathbf{w}}(t) = \mathbf{q}(t), \tag{1}$$

where $\mathbf{p}(t)$ and $\ddot{\mathbf{w}}(t)$ are the vectors of pressure differences and cochlear partition accelerations, $\mathbf{F}$ is the finite-difference matrix, and $\mathbf{q}(t)$ is the vector of source terms. The cochlear micromechanics of isolated partition segments are described by individual matrices. When Eq. (1) is substituted into an equation combining all the uncoupled elemental matrices, the coupled model of the cochlea can be described by the state space equations

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t), \tag{2}$$

and

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t), \tag{3}$$

where $\mathbf{A}$ is the system matrix, $\mathbf{x}(t)$ is the vector of state variables, $\mathbf{B}$ is the input matrix, $\mathbf{u}(t)$ is a vector of inputs proportional to $\mathbf{q}(t)$, $\mathbf{y}(t)$ is the output variable (BM displacement in this case), $\mathbf{C}$ is the output matrix, and $\mathbf{D}$ is an empty feedthrough matrix.

Figure 2 illustrates typical BM displacement responses to tonal stimuli. The phase lag of these responses at CF is similar to measurements made in the middle of the squirrel monkey cochlea (Robles and Ruggero, 2001).

The magnitude of the impedance mismatch between the interface of the middle ear and the cochlea can now also be determined. The nominal value of the cochlear model's characteristic impedance, Z_c, has been determined to be 2

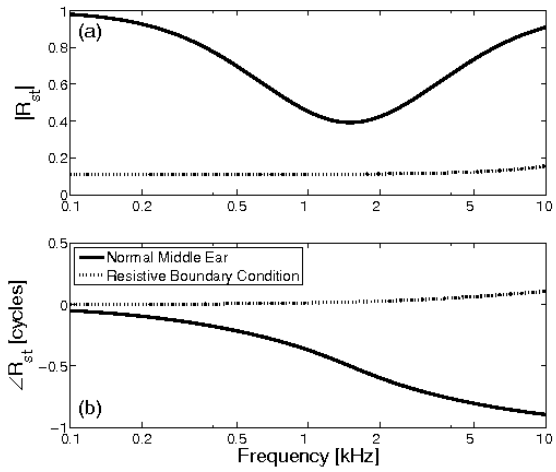


FIG. 3. Magnitude (a) and phase (b) of the basal reflection coefficient, R_{st}, given the base line middle ear (solid) and a largely resistive boundary (dotted).

$\times 10^{10}$ SI acoustic ohms. The reflection coefficient due to the middle ear as viewed from the cochlea, R_{st}, is given by Shera and Zweig (1990):

$$R_{st} = \frac{Z_{out} - Z_c}{Z_{out} + Z_c}. \tag{4}$$

The magnitude and phase of the state space model's reflection coefficient are plotted in Fig. 3 for the base line middle ear boundary, and also a resistance-dominated boundary ($C_{ME} = 8 \times 10^4$ N s m^{-3}).

Figure 4 shows the stability plot of the base line cochlear model given a nominal value of micromechanical feedback gain at all positions, $\gamma(x) = 1$. The stability plot shows the real (σ) and imaginary ($2\pi f$) parts of each of the poles of the coupled system, which are calculated from the eigenvalues of the system matrix, **A**, in Eq. (2). The imaginary components of the poles are converted from rad/s to kHz. For uniform values of feedback gain across the cochlea, the system becomes unstable at $\gamma(x) = 1.14$. This is indicated by the existence of at least one pole with a positive divergence rate, $\sigma(s) > 0$.

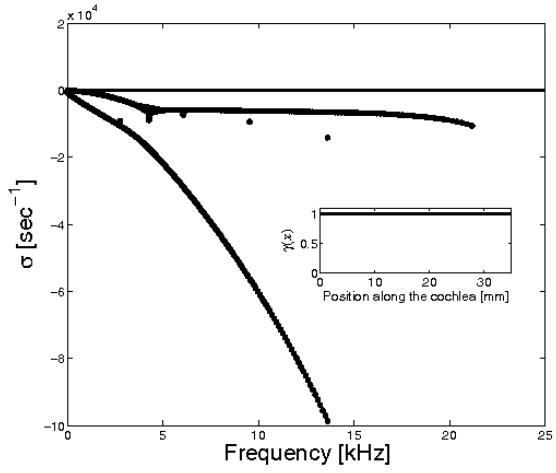


FIG. 4. A stability plot of the cochlear model given nominal gain, $\gamma(x) = 1$, and base line middle ear boundary.

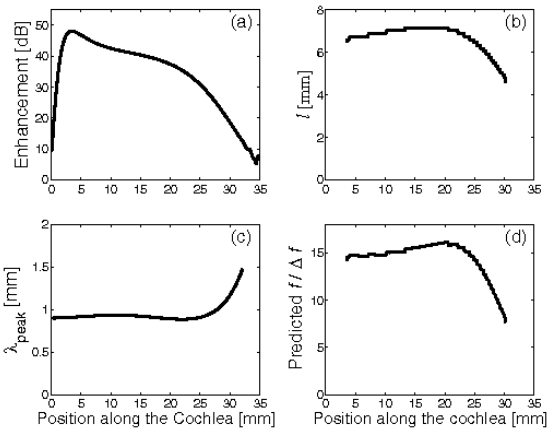


FIG. 5. Calculated characteristics of the model as a function of position along the cochlea: (a) active enhancement; (b) length scale, i.e., distance along the cochlea by which the characteristic frequency changes by a factor of e, as directly measured from the model; (c) λ_{peak}, wavelength of the TW in its peak region; (d) predicted spacing of SOAEs. Note that it was not possible to accurately calculate the length scale near the base and apex, hence the shortening of (b) and (d).

Shera (2003) argued that the CA is analogous to a laser's gain medium. One would expect a higher level of gain in the CA to result in greater system instability, given the same pattern of inhomogeneities in the cochlea. A higher value of feedback gain, $\gamma(x)$, results in greater active enhancement, which is defined here as the ratio of the cochlea's maximum active [$\gamma(x) = 1$] BM velocity to its maximum passive [$\gamma(x) = 0$] BM velocity across frequency at a given position, in dB. In the current model, the active enhancement is a function of position in the cochlea that is greatest (approximately 45 dB) near the base and gradually decreases toward the apex, shown in Fig. 5(a). This trend was demonstrated by Robles and Ruggero (2001), who made physiological measurements in animals.

According to Shera and Zweig (1993), the average distance between resonant positions of SOAEs along the cochlea is

$$\overline{\Delta x}_{SOAE} \approx \frac{1}{2}\lambda_{peak}, \tag{5}$$

where λ_{peak} is the wavelength of the TW in its peak region. Consequently, the predicted normalized spacing between SOAE frequencies is

$$f/\Delta f \approx 2l/\lambda_{peak}, \tag{6}$$

where l is the cochlear length scale, the distance over which the best frequency changes by a factor of e, shown for the model in Fig. 5(b). The normalized spacing is defined as the ratio of the geometric mean of two adjacent SOAE frequencies, f_a and f_b, divided by their difference,

$$f/\Delta f = \frac{\sqrt{f_a f_b}}{|f_a - f_b|}. \quad (7)$$

The PMD in humans is approximately 15 when expressed in terms of $f/\Delta f$ (Shera, 2003).

In order to calculate the wavelength of the TW for a given position and frequency in the state space model, it is necessary to return to the wave equation (de Boer, 1996):

$$\frac{\partial^2 p(x,\omega)}{\partial x^2} + \kappa_{\mathrm{TW}}^2(x,\omega)p(x,\omega) = 0, \quad (8)$$

where p is the pressure across the BM and κ_{TW} is the wave number of the TW, both functions of position and frequency. The wave number is related to the cochlear partition impedance, Z_{cp}, by the following:

$$\kappa_{\mathrm{TW}}^2(x,\omega) = \frac{-2j\omega\rho}{HZ_{\mathrm{cp}}(x,\omega)}, \quad (9)$$

where ρ is the density of the fluid, and H is the height of the scala vestibule and scala tympani above and below the cochlear partition. By definition,

$$\mathrm{Re}(\kappa_{\mathrm{TW}}) = \frac{2\pi}{\lambda_{\mathrm{TW}}}, \quad (10)$$

where λ_{TW} is the wavelength of the TW.

It is now possible to relate the wavelength of the TW in its peak region to the cochlear partition impedance at a given place, x, with characteristic frequency, ω_{cf},

$$\lambda_{\mathrm{peak}}(x) = \mathrm{Re}\left[\sqrt{\frac{HZ_{\mathrm{cp}}(x,\omega_{\mathrm{cf}})}{-2j\omega_{\mathrm{cf}}\rho}}\right]2\pi. \quad (11)$$

This is shown in Fig. 5(c), and is approximately 0.9 mm across most of the cochlea and slowly increases near the apex, thus breaking scaling symmetry. This trend is also consistent with physiological measurements made at the base and apex in animals (Robles and Ruggero, 2001).

Given the cochlear length scale and the wavelength of the TW at its peak as a function of position, the predicted spacing between unstable frequencies, $f/\Delta f$, can now be calculated as in Eq. (7). This result is shown in Fig. 5(d). The model's predicted SOAE spacing is approximately the measured PMD in humans ($f/\Delta f \approx 15$) for most of the length of the cochlea and decreases toward the apex.

III. SPATIALLY VARYING GAIN

It has been previously reported that deviations from a smoothly varying set of micromechanical parameters can cause instability in cochlear models. It is believed that the frequencies of cochlear instability represent the frequencies of potential SOAEs. Elliott *et al.* (2007) demonstrated that these models are most sensitive to rapid changes in the gain as a function of position. In the current paper, greater consideration is given to the nature of the inhomogeneities introduced and the resultant characteristics of the unstable frequencies. The feedback gain as a function of position along the cochlea, $\gamma(x)$, has been chosen as the parameter to be perturbed. In order to compare the relative level of instability present in a cochlea, it is instructive to examine the number of unstable frequencies present. However, to further quantify the magnitude of a cochlear model's instability, the concept of a pole's damping ratio is reviewed.

A second-order system can be described by its damping ratio, ζ, a dimensionless quantity that describes the rate at which system oscillations decay following an initial perturbation. This is related to the poles of a system, $s = \sigma + j\omega$, in the following manner:

$$\zeta = \cos(\alpha) = \frac{-\sigma}{\sqrt{\sigma^2 + \omega^2}}, \quad (12)$$

where α is the angle formed between the positive-real half-axis of the s-plane and the pole in question. When poles with nonzero imaginary components cross into the positive-real half-plane [$\zeta(s) < 0$], the response of a linear system will diverge exponentially. The rate of this divergence is given by $e^{-\zeta\omega_n t}$, where t is the time and ω_n is the resonant frequency of the pole in units of angular frequency. ω_n is determined by calculating the imaginary component of the pole. The damping ratio of an unstable pole is useful as it relates the rate at which the system will become unstable; the average value of many poles can also be compared across different cochlear models. This quantity is referred to as the *undamping ratio* in this paper, in the context of discussing unstable poles, and is assigned as ξ:

$$\xi = -\zeta. \quad (13)$$

A step change in gain is employed as a starting point for the discussion of cochlear stability analysis. From there, sinusoidal spatial variations and the band-limited random spatial variations are applied as gain distributions. It is important to note that the step and sinusoidal distributions of $\gamma(x)$ are introduced to understand the underlying mechanisms of SOAE generation and should not be interpreted as an attempt to model what necessarily exists in a human cochlea.

A. Step change in gain

A step change in gain gives rise to a discontinuity in the variation of BM impedance as a function of position along the cochlea. An ideal step in space has a well-distributed wave number spectrum, and thus should reflect wavelets across a wide range of wavelengths. One additional consequence of varying the gain as a function of position, $\gamma(x)$, is that the underlying properties of the TW are affected. For instance, a higher gain results in a shorter λ_{peak}. To minimize this effect, a relatively small amplitude step was chosen with a $\pm 3\%$ deviation from nominal gain on either side of the step. The stability plot for the cochlear model with such a step imposed on the gain at 18.2 mm from the base of the cochlea is shown in Fig. 6.

Three distinct frequencies are found to be unstable in this cochlea, at 1.478, 1.577, and 1.669 kHz. These frequencies are all close to the characteristic frequency at the location of the discontinuity, which is 1.550 kHz. According to Zweig and Shera (1995), only the frequencies whose responses peak in this region may become unstable since not enough energy is reflected otherwise; this is seen in Fig. 6 as

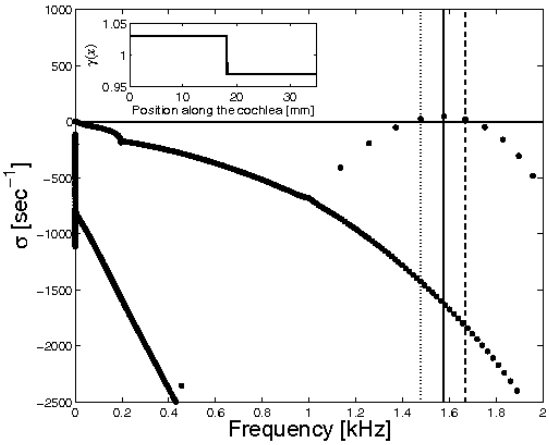


FIG. 6. Stability plot for a cochlea with the stepped gain inset: $\gamma(x<18.2\ \text{mm})=1.03$ and $\gamma(x\geq 18.2\ \text{mm})=0.97$. Note the frequency scale has been shorted to emphasize the locations of the unstable poles. Vertical lines are the frequencies of the unstable poles: 1.478 kHz (dotted), 1.577 kHz (solid), and 1.669 kHz (dashed).

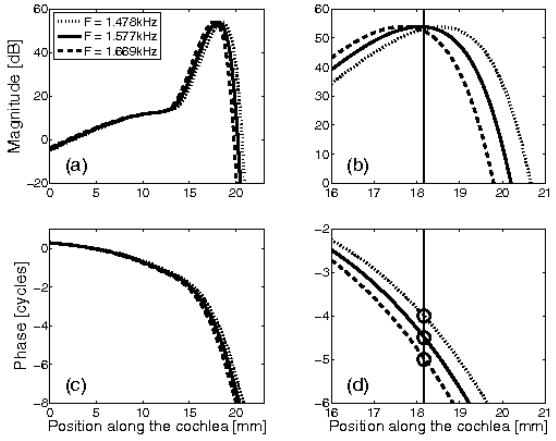


FIG. 7. Magnitude [(a) and (b)] and phase [(c) and (d)] of basilar membrane velocity for excitation at 1.478 kHz (dotted), 1.577 kHz (solid) and 1.669 kHz (dashed) given a base line model with nominal gain, $\gamma(x)=1$. (b) and (d) show the expanded axes for clarity of interpretation. A vertical line is drawn at the location of the discontinuity of Fig. 6 in the zoomed-in panels [(b) and (d)]. Circles in the phase plot (d) indicate phase shifts of −4, −4.5, and −5 cycles at this location.

only three frequencies near the discontinuity's characteristic frequency are unstable. Furthermore, there is a range of successively more stable poles that follow an arc leading away from the three unstable poles, both higher and lower in frequency. Presumably, the TWs of these frequencies are not reflected strongly enough by the discontinuity to cause instability.

The resultant spacings between the two pairs of adjacent unstable frequencies, $f/\Delta f$, are approximately 15 for the pair lower in frequency, and approximately 17 for the pair higher in frequency. This is consistent with the expectations given a slightly lower γ value apical of the discontinuity, and a slightly higher γ value basal to the discontinuity. To better understand why only these specific frequencies become unstable, Fig. 7 shows the magnitudes and phases of the BM velocity responses at these frequencies, for which a nominal gain throughout the cochlea is used, $\gamma(x)=1$.

A vertical line through Fig. 7(b) and Fig. 7(d) denotes the location along the cochlea of the discontinuity applied in Fig. 6. This line intersects with the phase responses of the 1.478, 1.577, and 1.667 kHz stimulus tones at −4, −4.5, and −5 cycles, respectively, within an accuracy of 1%. This is consistent with the "cochlear laser" theory of SOAE generation which requires that the phases of the unstable frequencies must undergo an integer number of cycles of total phase change between the reflection site and the middle ear boundary in order to combine constructively over successive reflections. For the unstable frequencies shown above, the "round-trip" phase change would equal 8, 9, and 10 cycles. Reexamining Fig. 6 in light of this feature, the stable poles that follow the same arc as the unstable poles must also represent frequencies that scatter wavelets which constructively combine, but perhaps are too weak to overcome the damping basal to the inhomogeneity.

Shera and Zweig (1993) and Zweig and Shera's (1995) concept of SOAE generation assumes wave amplification and multiple reflections between the middle ear boundary and the region of backscattering. A simple test of this theory involves changing the middle ear boundary so that it is less reflective.

Figure 8. shows the stability plot of a cochlear model with the same step change introduced in Fig. 6, but with a resistive boundary in the place of the human middle ear boundary, as shown in Fig. 3. The imaginary parts of the poles of Fig. 8 are almost identical to those of Fig. 6, but the real parts of the poles affected by the discontinuity are more stable. Whereas the base line model with a step change in gain was unstable, the model with the revised boundary and the same discontinuity is now stable.

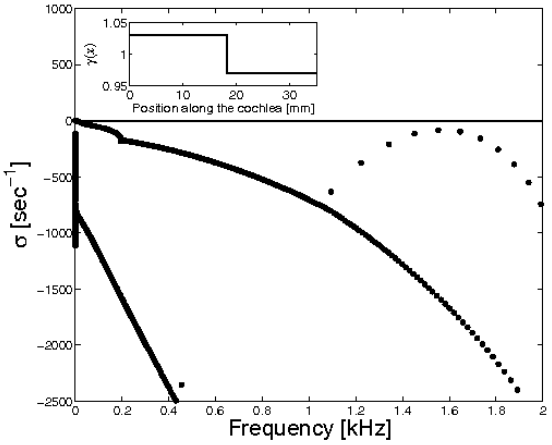


FIG. 8. Stability plot of a cochlear model with stepped gain distribution inset: $\gamma(x<18.2\ \text{mm})=1.03$ and $\gamma(x\geq 18.2\ \text{mm})=0.97$. The base line middle ear has been replaced with a resistive boundary, the reflection coefficient of which is shown in Fig. 3.

Appendix C

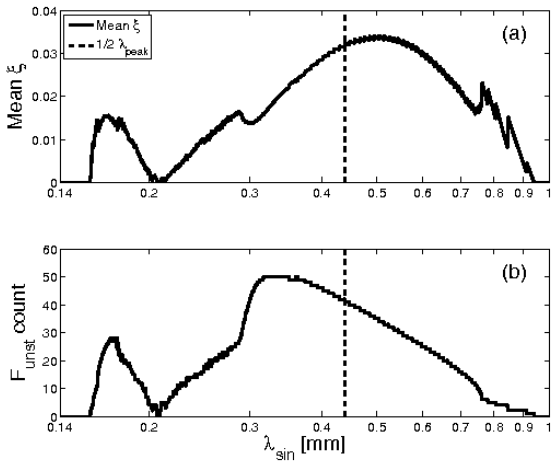


FIG. 9. Average undamping ratio ξ (a) and number of unstable frequencies (b) for a sinusoidal distribution of gain with varying wavelength, λ_{sin}. A vertical line marks the location of half the wavelength of the TW at its peak.

B. Sinusoidal variations in gain

A distribution of gain that is sinusoidal as a function of position is of interest because its wave number spectrum is concentrated at a single wave number, just as a sinusoidal wave form that is a function time has a frequency spectrum that is concentrated at a single frequency. This set of simulations follows the theory outlined by Strube (1989), which assumes uniform corrugations in gain along the BM. A range of wavelengths was chosen for the sinusoidal variation gain as a function of position along the cochlea, varying from 1 mm down to 0.14 mm, the latter being the spatial Nyquist limit of the model. A 10% peak-to-peak variation in amplitude about nominal gain generated instabilities over most of this spatial range, while maintaining stability for sinusoidal wavelengths greater than approximately 0.95 mm.

Figure 9 summarizes the level of instability in these cochleae by plotting both the mean undamping ratio, ξ, and the number of unstable frequencies as a function of the gain's sinusoidal wavelength. As expected, given the theories of Strube (1989), Shera and Zweig (1993), and Zweig and Shera (1995), the strongest instability occurred when the wavelength of the sinusoid, λ_{sin}, was approximately half the peak wavelength; this value occurs at 0.44 mm in the model. In addition, there was a region of greatly decreased instability in the model, centered about a periodicity of approximately one-fourth peak wavelength. This is thought to be due to destructive interference between the reflection sites, as the backscattered wavelets are out of phase with each other given this spatial periodicity.

The locally jagged aspect of the mean undamping ratio curve in Fig. 9(a) at approximately 0.75 mm is due to the periodic introduction of "new" unstable poles with low-undamping ratios that are generated as the wavelength of the sinusoid is varied. The average undamping ratio peaks at approximately 0.5 mm, which is slightly longer than half the peak wavelength for most of the length of the cochlea in this model. It is of note that the maximum in the total number of unstable poles, shown in Fig. 9(b), is located at a sinusoidal wavelength somewhat shorter than half the peak wavelength. As the sinusoidal wavelength of the gain variations is shortened, the number of peaks in the gain (and thus reflection sites) along the cochlea increases, creating more unstable poles. Even for sinusoidal periods less than half the peak wavelength, the rate at which unstable poles are being generated per unit decrease in λ_{sin} is still outpacing the rate at which they are returning to stability; this explains the location of the peak in Fig. 9(b).

C. Band-limited random gain distributions

Shera and Zweig's (1993) theory of SOAE generation assumes that the cochleae of normal-hearing humans contain a dense but random array of inhomogeneities. Each of these place-fixed perturbations reflects energy from the forward TW (Talmadge *et al.*, 1993; Shera and Zweig, 1993; Zweig and Shera, 1995). In this section, the stability of cochlear models with band-limited, spatially random gain distributions is used to approximate what is postulated to exist in a human cochlea. A fifth order Butterworth filter was employed to band-limit gain distributions in the wave number domain (Lineton, 2001). The low wave number cutoff frequency was fixed at the length of the cochlea itself, in order to prevent any dc shifts in the gain. The high wave number cutoff frequency was initially set to 6.6 radians/mm and slowly increased, thus generating cochlear models with successively more densely spaced reflection sites. The average filter bandwidths have been plotted below in terms of 2π times inverse wave number; this quantity has units of length (mm) and is directly comparable to the wavelength of the TW at its peak.

Figure 10 summarizes the results of simulations of 400 different cochleae, each with unique, spatially random gain distributions. The (a) panels show a typical stability plot from each group. The averaged power spectrum of the gain distributions is shown in the (b) panels, the two Roman numeral sets (I and II) having different high wave number (low wavelength) cutoffs; the dashed vertical line represents half the wavelength of the TW at its peak. A 5 mm sample of a gain distribution at this cutoff wavelength is inset. The (c) panels show the histograms of the average number of unstable poles per cochlea, sorted into logarithmic frequency bins. Figure 10(IIc) demonstrates that a lower cutoff wavelength, and thus a more rapid spatial variation in gain, is necessary to generate instability at frequencies below 2 kHz. This is believed to be due to the lower level of enhancement toward the apex of the cochlea and the lower magnitude of the basal reflection coefficient in this frequency band.

The histogram of normalized spacings of adjacent unstable frequencies per cochlea is shown in the (d) panels. The data for the (c) and (d) panels are presented for all instabilities (gray, thick bars) and also in a restricted range of 0.5–2 kHz (thin, black bars). This smaller range represents the frequency bandwidth where the middle ear's reverse-pressure transfer function is most efficient (Puria, 2003), and thus where one might expect the most SOAEs to be detected. The results for λ_{cutoff}=0.19 mm [Fig. 10(IId)] show a peak in

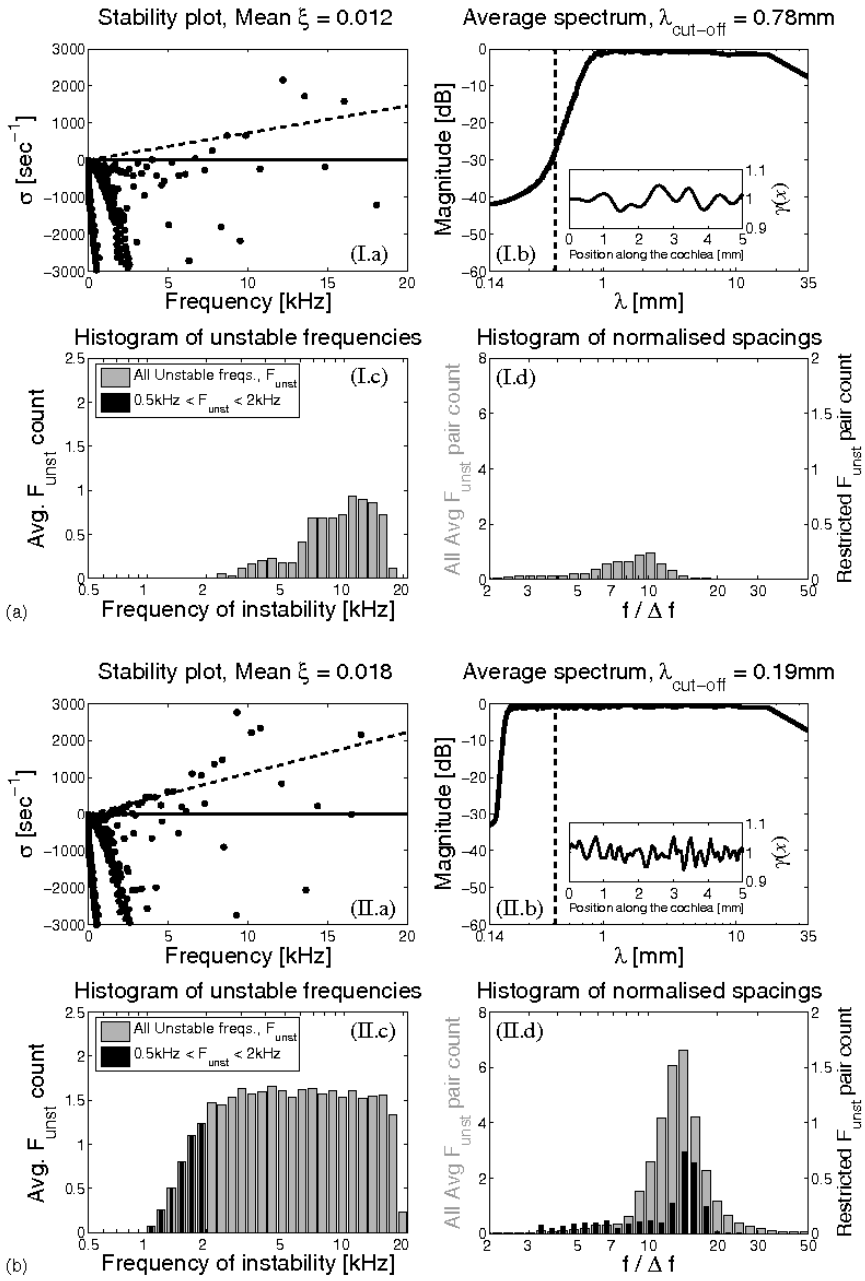


FIG. 10. The collected results from 2×200 cochlear models with randomly generated gain distributions. Each Roman numeral subset has been filtered with a different cutoff wavelength: [(Ia)–(Id)] λ_{cutoff}=0.78 mm, [(IIa)–(IId)] λ_{cutoff}=0.19 mm. A peak-to-peak amplitude of 15% was applied to these gain distributions. (a) A characteristic stability plot taken from the set. The average undamping ratio for that single case, ξ, is given and superimposed (dotted line). (b) Averaged inverse wave number spectrum of the gain distributions. Half the peak wavelength is indicated by a dotted vertical line and the first 5 mm of a characteristic gain distribution are inset. (c) Averaged histogram of all unstable frequencies per cochlea sorted in logarithmic frequency bins (gray). The instabilities occurring between 0.5 and 2 kHz are superimposed in thin, black bars. (d) Averaged histogram of normalized spacings ($f/\Delta f$) per cochlea in gray. The histogram of spacings of unstable frequencies per cochlea occurring between 0.5 and 2 kHz is again superimposed in black. Note the different left- and right-vertical scales.

the normalized spacing at $f/\Delta f \approx 15$ in the frequency range of 0.5–2 kHz. These results are consistent with the Shera and Zweig's (1993) theory which assumes a dense array of reflection sites, represented in these simulations by a low cutoff wavelength.

Figure 11 summarizes data from the above calculations, while also presenting data from many other simulations which have different cutoff wavelengths and peak-to-peak variations in gain. The mean unstable frequency count and the mean undamping ratio, ξ, vary directly with the ampli-

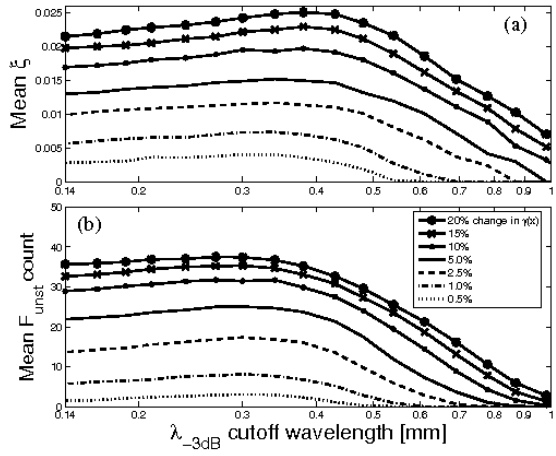


FIG. 11. Variation of average undamping ratio (a) and total unstable pole count (b) with cutoff wavelength for five amplitudes of peak-to-peak random variations in gain. At each amplitude, 20 cutoff wavelengths were each applied to 200 models with randomly generated gain distributions. A total of 28,000 stability tests were performed.

tude of the variation in $\gamma(x)$. This result is consistent with the findings of Elliott *et al.* (2007). In contrast to the sinusoidal case (see Fig. 9), no distinct notch in either the average undamping ratio or the number of instabilities is apparent at a cutoff wavelength of approximately one-quarter of the peak wavelength. In the sinusoidal simulations, all of the spatial spectral energy was concentrated at a particular wave number; this potentially generated strong, destructive interference when the sinusoidal wavelength was one-quarter the peak wavelength. The spectral energy in the random spatial variations in gain is comparatively much more diffuse, perhaps reducing the amount of both constructive and destructive interferences. The statistics of the spacings of instabilities is thus largely independent of the exact form of the spatial variations, provided they have a significant component at the wave number corresponding to one-half λ_{peak}. Peak-to-peak variations in $\gamma(x)$ as small as 0.5% can give rise to instabilities, provided λ_{cutoff} is less than approximately 0.5 mm, near the half peak wavelength.

IV. DISCUSSION

The findings of this paper, based on a numerical model of the human cochlea, are consistent with the multiple-reflection theory of Zweig and Shera (1995). The state space formulation is able to predict the frequencies at which a linear, active cochlear model will become unstable. Elliott *et al.* (2007) presented a nonlinear time-domain simulation demonstrating that an isolated unstable pole will evolve into a limit cycle within the cochlea at the expected frequency. Direct measurements in animals have shown that spontaneous basilar membrane oscillation is associated with SOAEs in the ear canal (Nuttall *et al.*, 2004). Consequently, comparisons are drawn in this paper between measured SOAE characteristics and the instabilities generated in the cochlear model. However, it is worth highlighting the similarities and differences between measured data and these simulation results.

This model predicts that instabilities exist all along the cochlea and across a wide band of frequencies, given a dense array of inhomogeneities in the cochlea. In contrast, SOAEs in normal-hearing individuals are only routinely detected between 0.5 kHz and 6 kHz (Probst *et al.*, 1991). Even if instabilities exist in all regions along the average human cochlea, however, it is likely that only a subset of these will be detected in the ear canal. It is believed that the inefficient reverse-transmission characteristics of the middle ear hinder the detection of SOAEs outside of its best transmissibility range, given its steep drop-off below and above resonance, of approximately −40 dB/decade. The limited bandwidth of normally detected SOAEs is also potentially reduced by physiological noise and the current limitations of sensor technology. Just as improved measurement techniques have revealed increasingly sharp active BM enhancement through the years, refinements in recording technique have exposed a higher prevalence of SOAEs in more recent studies (Probst *et al.*, 1991; Penner and Zhang, 1997).

The average number of unstable frequencies shown in Fig. 11 for a 10% peak-to-peak variation in gain is similar to the maximum number of emissions detected in a single ear, some in excess of 30 SOAEs (Talmadge *et al.*, 1993). It has been shown that the level and number of instabilities in the state space model depend on the amplitude of the variations in BM impedance and the spatial density of the inhomogeneities. When nonlinear effects are incorporated into time domain simulations, it is anticipated that the total number of instabilities may differ from those predicted by linear stability analysis.

It has been demonstrated by numerous experimentalists (e.g., Zwicker and Schloth, 1984) that externally applied stimuli can frequency-lock, phase-synchronize, suppress, or otherwise affect a SOAE. Some modelers have used Van der Pol oscillators to account for these phenomena (Bialek and Wit, 1984; Wit, 1986; van Hengel *et al.*, 1996). Further work is needed to examine the nonlinear interaction of limit cycles and external stimuli in the state space model presented here.

The current linear model predicts a distribution of unstable frequency spacings that is similar to physiologically compiled data in several respects. Although the current model's results do not accurately match the observed variation in SOAE spacings with frequency (Shera, 2003), the spacing results presented in Fig. 10(IId) are consistent with the predictions shown in Fig. 5(d). Furthermore, the peak in the normalized spacings of Fig. 10(IId) is correctly located at the PMD when sufficient spectral content is present in the inhomogeneities at half the peak wavelength, as predicted by Zweig and Shera (1995).

When the current understanding regarding hearing sensitivity, the various forms of OAEs and pathology are combined, a convincing picture regarding the generation of SOAEs begins to emerge. As many authors have pointed out, SOAEs in humans appear to be a natural by-product of the species' sharply tuned sense of hearing. Normal hearing individuals that do not exhibit SOAEs typically have an audio-

gram which underperforms those with SOAEs by approximately 3 dB in the standard 1–6 kHz range (McFadden and Mishra, 1993). Pélanová *et al.* (2007) also reported that the high-frequency audiogram of normal-hearing children without SOAEs underperformed those with SOAEs by approximately 5 dB through the 10–16 kHz range. In the "laser-cochlea" theory of OAE generation, it is the portion of the cochlea basal to the reflection site that is crucial to sustaining the limit cycle oscillation. If the losses in this region are not overcome by the active enhancement provided by the outer hair cells, no spontaneous emission can occur.

V. CONCLUSIONS

Simulations using the state space model of the human cochlea show patterns of SOAE production that can be explained by Zweig and Shera's (1995) theory. As demonstrated by the step change in gain, only frequencies with a TW that undergoes an integer round-trip phase change between the middle ear boundary and the inhomogeneity will become unstable. Instabilities are detected along the entire cochlea given spatially random changes in gain, but it is believed that only a subset of these unstable frequencies become measurable as SOAEs due to the middle ear's inefficient reverse transmission characteristics. The spectral content of the inhomogeneities in the BM impedance also has a strong impact upon the level and frequency spacings of the resultant instabilities.

A 10% variation in gain as a function of position generated the most instability in the model when a sinusoidal inhomogeneity with a wavelength roughly equal to half the wavelength of the TW at its peak was applied; instability was eliminated when the sinusoid's wavelength was reduced to roughly one-fourth the wavelength of the TW at its peak. When random inhomogeneities are simulated, the expected PMD between adjacent unstable frequencies is strongly expressed in the results only when there is sufficient spectral content at one-half the wavelength of the TW at its peak.

Nonlinear time-domain simulations, such as those introduced in Elliott *et al.* (2007), are expected to provide a method of explaining the more subtle interactions that exist in human cochleae due to multiple instabilities and externally applied stimuli. However, it is clear that this linear model can provide a great deal of insight into the mechanisms underlying the generation of SOAEs as numerical results presented here are in good agreement with the theory of Zweig and Shera (1995).

ACKNOWLEDGMENTS

The authors would like to thank Dr. Sunil Puria for sharing his research data, and two anonymous reviewers for their helpful comments and suggestions. This work was partially supported by a Fulbright Postgraduate Award.

APPENDIX: HELICOTREMA BOUNDARY CONDITION

The structure of the noncoupled micromechanical matrices is identical to Elliott *et al.* (2007), except at the helicotrema for which

$$\dot{\mathbf{x}}_N(t) = \mathbf{A}_N\mathbf{x}_N(t) + \mathbf{B}_N p_N(t), \tag{A1}$$

where the boundary condition is now taken to be a mass-damper system, so that

$$\mathbf{x}_N(t) = [\dot{w}'_N(t) \quad w'_N(t)]^T, \tag{A2}$$

$$\mathbf{A}_N = \begin{bmatrix} \dfrac{-C_H}{M_H} & 0 \\ 1 & 0 \end{bmatrix}, \tag{A3}$$

and

$$\mathbf{B}_N = \begin{bmatrix} \dfrac{1}{M_H} & 0 \end{bmatrix}^T. \tag{A4}$$

In order to incorporate this change into the macromechanical formulation, it was necessary to insert an additional term in the finite difference fluid-coupling matrix, **F**, such that it is still invertible. The expanded matrices represented in Eq. (1) of this work now become

$$\frac{H}{2\rho\Delta^2}\begin{bmatrix} -\dfrac{\Delta}{H} & \dfrac{\Delta}{H} & & & & & & 0 \\ 1 & -2 & 1 & & & & & \\ 0 & 1 & -2 & 1 & & & & \\ & & \ddots & \ddots & \ddots & & & \\ & & & 1 & -2 & 1 & & 0 \\ & & & & 1 & -2 & & 1 \\ 0 & & & & & \dfrac{\Delta}{H} & & -\left(\dfrac{\Delta}{H}+\dfrac{\Delta^2}{H^2}\right) \end{bmatrix} \times \begin{bmatrix} p_1(t) \\ p_2(t) \\ \vdots \\ p_{N-1}(t) \\ p_N(t) \end{bmatrix} - \begin{bmatrix} \ddot{w}_{SR}(t) \\ \ddot{w}_2(t) \\ \vdots \\ \ddot{w}_{N-1}(t) \\ \ddot{w}'_N(t) \end{bmatrix} = \begin{bmatrix} \ddot{w}_{SO}(t) \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \tag{A5}$$

where H is the height of the channel, ρ is the density of the fluid, and Δ is the length of a cochlear segment. The physical meaning of this additional term in the fluid-coupling matrix can be determined by relating this revised equation to the boundary condition at the apex.

The last row in Eq. (A5) represents the helicotrema boundary condition and can be written as

$$\frac{H}{\Delta^2}\left[\frac{\Delta}{H}p_{N-1} - \left(\frac{\Delta}{H}+\frac{\Delta^2}{H^2}\right)p_N\right] = 2\rho\ddot{w}'_N, \tag{A6}$$

where p_{N-1} and p_N are the pressures adjacent to and at the helicotrema, and $\ddot{w}'_N$ represents the "effective" helicotrema acceleration. Rewriting Eq. (A6) gives

$$\frac{p_{N-1}-p_N}{\Delta} = 2\rho\ddot{w}'_N + \frac{1}{H}p_N. \tag{A7}$$

However, the physical boundary condition is defined as

Appendix C

$$\left.\frac{dp}{dx}\right|_{x=L} = 2\rho\ddot{w}_N, \tag{A8}$$

where $\ddot{w}_N$ is the true helicotrema acceleration. Using a finite difference approximation in Eq. (A8) and expressing the acceleration at the helicotrema as a velocity yields

$$\frac{p_{N-1} - p_N}{\Delta} \approx 2j\omega\rho\dot{w}_N. \tag{A9}$$

The true admittance at the helicotrema is the volume velocity at the helicotrema divided by the local pressure:

$$Y_N = \frac{A\dot{w}_N}{p_N}, \tag{A10}$$

where A is the area of the helicotrema. Relating the approximated boundary condition in terms of the admittance gives

$$\frac{p_{N-1} - p_N}{\Delta} \approx \frac{2j\omega\rho}{A} Y_N p_N. \tag{A11}$$

The effective velocity at the helicotrema can also be expressed in terms of an effective admittance at the helicotrema, Y'_N [defined by the parameters in Eqs. (A1)–(A4)]:

$$\dot{w}'_H = \frac{Y'_N p_N}{A}. \tag{A12}$$

Substituting Eqs. (A11) and (A12) back into Eq. (A7) results in an equation that relates the true helicotrema admittance to its effective value:

$$\frac{2j\omega\rho}{A} Y_N p_N = \frac{2j\omega\rho}{A} Y'_N p_N + \frac{1}{H} p_N. \tag{A13}$$

Simplifying Eq. (A13) reveals

$$Y_N = Y'_N + \frac{A}{2j\omega\rho H}. \tag{A14}$$

The term $A/2j\omega\rho H$ is equivalent to the admittance of an acoustic mass of $m = 2\rho H/A$. The acoustic mass of a short tube of length L and area A is $2\rho L/A$. In this case, $H = L$ corresponds to the assumed length of the helicotrema opening. The assigned value of 1 mm corresponds well with the value quoted by Fletcher (1953). The added term in the finite difference fluid-coupling matrix can be interpreted as an inertial term in parallel with the effective helicotrema impedance which is defined by the state space model in Eqs. (A1)–(A4). It should be noted that the change to the helicotrema boundary condition has a negligible effect on the model's response above approximately 200 Hz. Below this frequency, the reflections from the apex are more strongly attenuated than when using the pressure release boundary condition presented in Elliott *et al.* (2007).

[1] Z_{out} is referred to as M3 in Puria (2003).

Allen, J. B., Shaw, G., and Kimberley, B. P. (**1995**). "Characterization of the nonlinear ear canal impedance at low sound levels," Assoc. Res. Otolaryngol. Abstr. **18**, 190.

Bialek, W., and Wit, H. (**1984**). "Quantum limits to oscillator stability: theory and experiments on acoustic emissions from the human ear," Phys. Lett. **104**, 173–178.

Braun, M. (**1997**). "Frequency spacing of multiple spontaneous otoacoustic emissions shows relation to critical bands: A large-scale cumulative study," Hear. Res. **114**, 197–203.

Dallmayr, C. (**1985**). "Spontane oto-akusticsche Emissionen: Statistik und Reaktion auf akustische Störtöne," Acustica **59**, 67–75.

Dallmayr, C. (**1986**). "Stationäre und dynamische Eigenschaften spontaner und simultan evozierter oto-akustischer Emissionen," dissertation, Technische Universität, Munich.

de Boer, E. (**1996**). "Mechanics of the cochlea: modeling efforts," in *The Cochlea*, edited by P. Dallos, A. N. Popper, and R. R. Fay (Springer-Verlag, New York), pp. 258–317.

Diependaal, R. J., de Boer, E., Viergever, M. A. (**1987**). "Cochlear power flux as an indicator of mechanical activity," J. Acoust. Soc. Am. **82**, 917–926.

Elliott, S. J., Ku, E. M., and Lineton, B. (**2007**). "A state space model for cochlear mechanics," J. Acoust. Soc. Am. **122**, 2759–2771.

Fletcher, H. (**1953**). in *Speech and Hearing in Communication*, edited by J. B. Allen (Acoustical Society of America, New York), pp. 248.

Greenwood, D. D. (**1990**). "A cochlear frequency-position function for several species–29 years later," J. Acoust. Soc. Am. **87**, 2592–2605.

Gold, T. (**1948**). "Hearing. II. The physical basis of the action of the cochlea," Proc. R. Soc. London, Ser. B **135**, 492–498.

Kemp, D. T. (**1979**). "Evidence of mechanical nonlinearity and frequency selective wave amplification in the cochlea," Arch. Oto-Rhino-Laryngol. **224**, 37–45.

Lineton, B. (**2001**). "Testing a model of the stimulus frequency otoacoustic emissions in humans," Ph.D. thesis, University of Southampton, Southampton.

Martin, P., and Hudspeth, A. J. (**1999**). "Active hair-bundle movements can amplify a hair cell's response to oscillatory mechanical stimuli," Proc. Natl. Acad. Sci. U.S.A. **96**, 14380–14385.

McFadden, D., and Mishra, R. (**1993**). "On the relation between hearing sensitivity and otoacoustic emissions," Hear. Res. **71**, 208–213.

Neely, S. T. (**1981**). "Finite difference solution of a two-dimensional mathematical model of the cochlea," J. Acoust. Soc. Am. **69**, 1386–1393.

Neely, S. T., and Kim, D. O. (**1986**). "A model for active elements in cochlear biomechanics," J. Acoust. Soc. Am. **79**, 1472–1480.

Nobili, R., Vetesnik, A., Turicchia, L., and Mammano, F. (**2003**). "Otoacoustic emissions from residual oscillations of the cochlear basilar membrane in a human ear model," J. Assoc. Res. Otolaryngol. **4**, 478–494.

Nuttall, A. L., Grosh, K., Zheng, J., de Boer, E., Zou, Y., and Ren, T. (**2004**). "Spontaneous basilar membrane oscillation and otoacoustic emission at 15 kHz in a guinea pig," J. Assoc. Res. Otolaryngol. **5**, 337–348.

Pélanová, J., Groh, D., Popear, J., Kabelka, Z., and Syka, J. (**2007**). "Presence and characteristics of spontaneous otoacoustic emissions in children and adolescents," Inner Ear Biology, Vol. **113**, p. 26.

Penner, M. J., and Zhang, T. (**1997**). "Prevelance of spontaneous otoacoustic emissions in adults revisted," Hear. Res. **103**, 28–34.

Probst, R., Lonsbury-Martin, B. L., and Martin, G. K. (**1990**). "A review of otoacoustic emissions," J. Acoust. Soc. Am. **89**, 2027–2066.

Puria, S. (**2003**). "Measurements of human middle ear forward and reverse acoustics: Implications for otoacoustic emissions," J. Acoust. Soc. Am. **113**, 2773–2789.

Robles, L., and Ruggero, M. A. (**2001**). "Mechanics of the mammalian cochlea," Physiol. Rev. **81**, 1305–1352.

Shera, C. A. (**2003**). "Mammalian spontaneous otoacoustic emissions are amplitude-stabilized cochlear standing waves," J. Acoust. Soc. Am. **114**, 244–262.

Shera, C. A., Tubis, A., and Talmadge, C. L. (**2003**). "Stimulus-spectrum irregularity and the generation of evoked and spontaneous otoacoustic emissions: comments on the model of Nobili *et al.*" Last viewed 12/20/2007. http://otoemissions.org/whitepapers/biophysics/chris_shera.html

Shera, C. A., and Guinan, J. J. (**1999**). "Evoked otoacoustic emissions arise by two fundamentally different mechanisms: A taxonomy for mammalian OAEs," J. Acoust. Soc. Am. **105**, 782–798.

Shera, C. A., and Zweig, G. (**1990**). "Reflection of retrograde waves within the cochlea and at the stapes," J. Acoust. Soc. Am. **89**, 1290–1305.

Shera, C. A., and Zweig, G. (**1993**). "Order from chaos: resolving the paradox of periodicity in evoked otoacoustic emission," *Biophysics of Hair Cell Sensory Systems*, edited by H. Duifhuis, J. W. Horst, P. van Dijk, and S. M. van Netten (World Scientific, Singapore), pp. 54–63.

Strube, H. W. (**1989**). "Evoked otoacoustic emissions as cochlear Bragg reflections," Hear. Res. **38**, 35–46.

Talmadge, C. L., Long, G. R., Murphy, W. J., and Tubis, A. (**1993**). "New

Appendix C

off-line method for detecting spontaneous otoacoustic emissions in human subjects," Hear. Res. **71**, 170–182.

Talmadge, C. L., and Tubis, A. (**1993**). "On modeling the connection between spontaneous and evoked otoacoustic emissions," in *Biophysics of Hair Cell Sensory Systems*, edited by H. Duifhuis, J. W. Horst, P. van Dijk, and S. M. van Netten (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**, 1517–1543.

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**, 3566–3571.

Wit, H. P. (**1986**). "Statistical properties of a strong spontaneous otoacoustic emission," in *Peripheral Auditory Mechanisms*, edited by J. B. Allen, J. L. Hall, A. E. Hubbard, S. T. Neely, and A. Tubis (Springer-Verlag, Berlin), pp. 221–228.

Zweig, G., and Shera, C. A. (**1995**). "The origin of periodicity in the spectrum of evoked otoacoustic emissions," J. Acoust. Soc. Am. **98**, 2018–2047.

Zwicker, E., and Peisl, W. (**1990**). "Cochlear preprocessing in analog models, in digital models, and in human inner ear," Hear. Res. **44**, 209–216.

Zwicker, E., and Schloth, E. (**1984**). "Interrelation of different oto-acoustic emissions," J. Acoust. Soc. Am. **75**, 1148–1154.

Appendix D

Growth Curves of CEOAEs

The growth of the RMS CEOAE amplitude is compared over different time windows in this appendix. Figure D.1.a-d shows the growth rates of the DNL- and DD- CEOAEs in short (a, b) and long (c, d) time windows. This is derived from the CEOAEs plotted in Figure 5.11and Figure 5.14.

As the stimulus amplitude is increased from linear to moderately saturating levels, the DD CEOAE increases at a rate that is slightly less than linear and the DNL CEOAE rises at a slope greater than unity. These results make sense, as the CA is only beginning to saturate. At moderate levels, starting around 40 dB SPL, the amplitude of the directly-determined CEOAE begins to level off and decrease. Again this matches expectations: as the CA reaches its heavily-saturating region, the variations in $\gamma(x)$ due to the inhomogeneities should become less significant due to the near-passive behaviour of the cochlea, thus reducing the level of reflections. What is contrary to predictions is the linear rise in CEOAE amplitude at stimulus levels > 72 dB SPL in the first ms (a, b), and > 96 dB SPL in the latter 20 ms (c, d). This may be due to reflections resulting from the passive mechanics of the cochlear model.

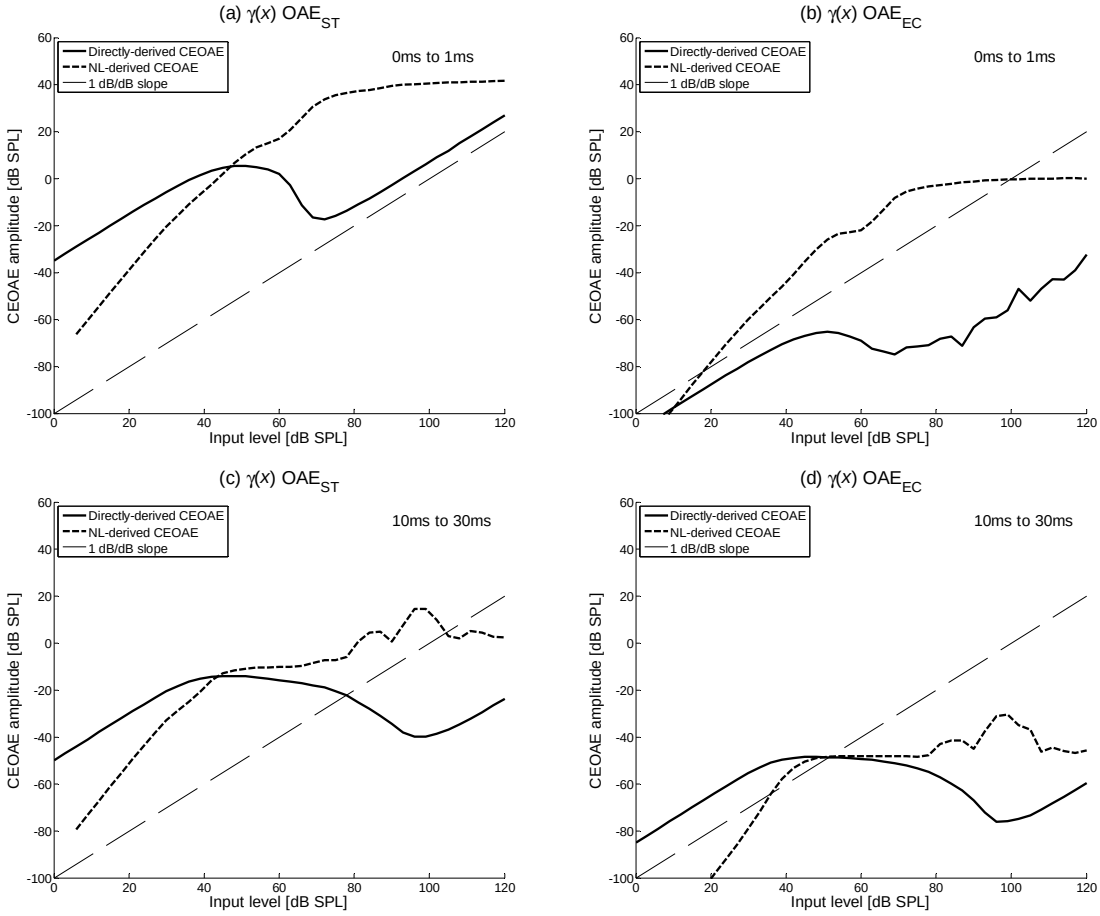


Figure D.1.a-d: Growth curves of CEOAE amplitude in (a, b) the first ms, and (c, d) the last 20 ms. Response at the stapes is shown in the left panels (a, c) and the response at the ear canal is shown on the right panels (b, d).

Appendix D

The variation of the δ(*x*)-perturbed CEOAE amplitude with level is shown in Figure D.2.a-b. These results are calculated from the results shown in Figure 5.16. Whereas the DD CEOAE amplitude increases at high levels in the model with the perturbation in γ(*x*), the equivalent plot here shows that the CEOAE amplitude reaches a near-constant value when the perturbation is located in δ(*x*). It is also significant that the slopes of the growth curves are approximately 2 dB/dB for stimulus levels < ~60 dB SPL for both the nonlinearly- and directly-determined curves.

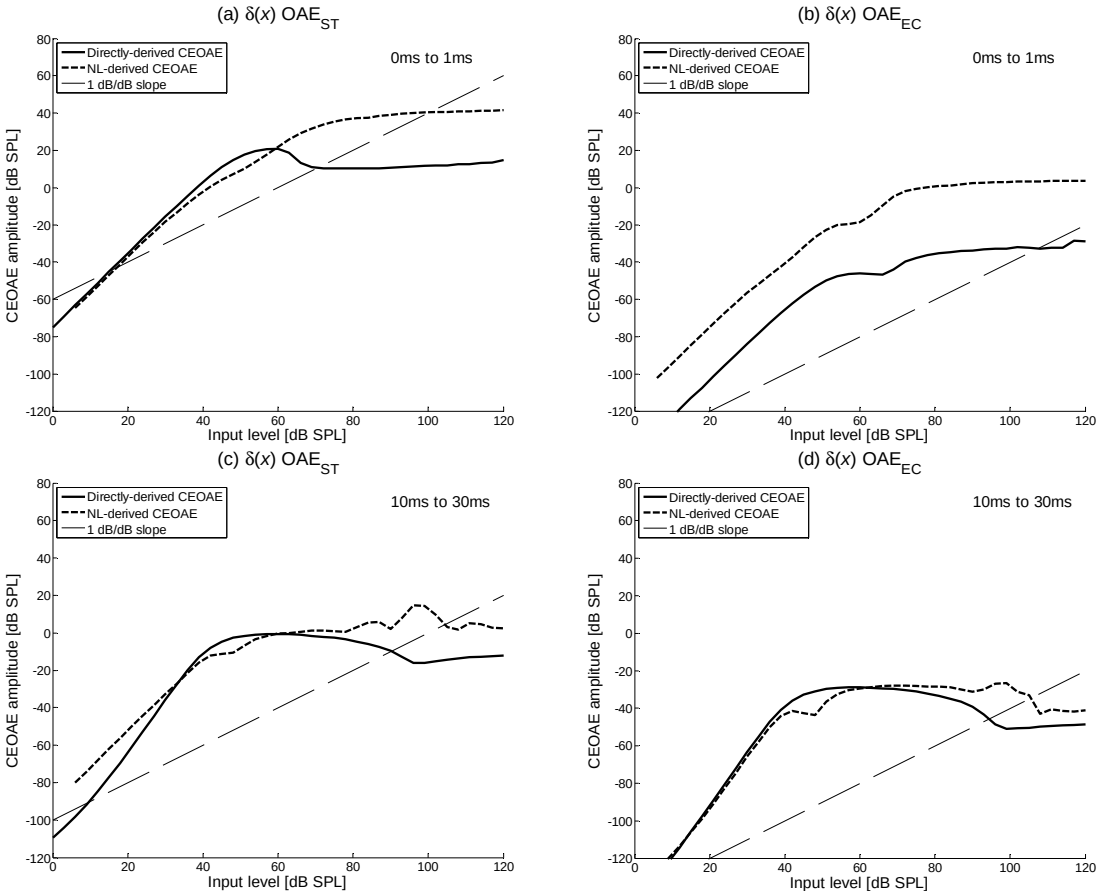


Figure D.2.a-b: Growth curves of CEOAE amplitude in short (a, b) and long (c, d) time windows at the stapes (a, c) and the ear canal (b, d). Perturbations were applied in δ(*x*).

The fact that the ESLE grows linearly given perturbations in γ(*x*) at high amplitudes suggests that a distortion explanation is unlikely; such phenomena are typically compressive in nature. Furthermore, that the ESLE waveform grows linearly in the model with inhomogeneities in γ(*x*) but not in the model with perturbations in δ(*x*) suggests that slightly different mechanisms may be operating here. This is still somewhat perplexing;

one would expect both forms of gain-based perturbation to become suppressed at the highest levels. Clearly more simulations of nonlinear CEOAEs should be run and ideally compared with measurements of the ESLE. One significant complication is that, due to its linear nature, traditional nonlinear methods of stimulus artefact cancellation also necessarily remove this component. For this reason, Kruglov *et al.* (1997) support use of linear methods. Another difficulty of determining the characteristics of the ESLE is that it is dominated by high frequency components in this model which are difficult to discern when transmitted through the middle ear.

The results presented here are further discussed in Chapter 6.

Appendix E

Suppression of Van der Pol Oscillator Simulations

This appendix describes some simulations of the suppression of Van der Pol Oscillators that were performed early in this Ph.D. Numerous authors have shown that the responses of SOAEs to external tones are very similar to a negatively damped, driven Van der Pol Oscillator (e.g., Murphy et al., 1995; van Dijk and Wit, 1990). This has encouraged some authors to suggest that the response of SOAEs in the ear canal can be directly related back to the local activity of such resonators along the BM. As a result, a number of cochlear models consisting of coupled Van der Pol Oscillators have been proposed (e.g., Sisto and Moleti, 1999). However, as discussed by Shera (2003), the success of the global standing wave model of SOAE generation (and the supporting simulations presented in this thesis) suggest that OAEs are due to the response of spatially distributed reflection and distortion sites throughout the cochlea.

The Van der Pol Oscillator nevertheless remains an interesting tool for studying the suppression of SOAEs in the cochlea, because it is able to encapsulate many features of the global response in a much-simplified form.

E.1 the Van der Pol Equation

The normalised equation for a driven Van der Pol Oscillator is given by:

$$\ddot{x} + \gamma\left(|x|^2 - 1\right)\dot{x} + \omega_n{}^2 x = A\cos(\omega_d t), \tag{E.1}$$

where x is the displacement, A is the driving amplitude, γ is the nonlinear damping term, and ω_n and ω_d are the natural and driven frequencies, respectively. This equation represents the standard formulation of a driven one degree-of-freedom simple harmonic oscillator, except that it includes a nonlinear damping term, i.e. the term proportional to $|x|^2$ inside the brackets in equation (E.1).

The following sections are a brief outline of the effect of each parameter in the equation; these ideas are further explored and illustrated in subsequent sections of this report.

Damping Factor γ

The viscous damping term in the Van der Pol equation consists of a nonlinear component and a linear component. The nonlinear term is always positive, and is proportional to the square of the response. When the linear damping term is negative, as in (E.1), the system injects energy into the motion of the mass within a range of displacements about the equilibrium position. If the linear damping term was set positive so that the damping term reads $\gamma(|x|^2+1)$, it would act as a normal viscous damper dissipating energy (and no longer reaches a limit cycle oscillation when not driven). In order to simplify analysis, the linear and nonlinear terms have been combined into a single coefficient, γ. For the purposes of this investigation, the value of γ was set to unity.

Driving Amplitude A

At zero and very low driving amplitudes, the system responds roughly as if it were an unforced oscillator, slowly ramping up to its natural limit cycle. In contrast, at high values of A, the system behaves as a relaxation oscillator. This is to say that periods of little motion are contrasted with short instances where the mass moves very quickly between the positive and negative extremes of its displacement. The values of A that lie between these two extremes are of greatest interest.

As the driving amplitude is increased the (steady-state) response quickly changes. Before a particular value of A, the steady-state response is dominated by the natural frequency of the system; beyond this point, the system oscillates at the driving frequency. This phenomenon is known as quenching of the limit cycle oscillation and the two regions are separated by a pull-out amplitude.

Appendix E

Driving Frequency Relative to Natural Frequency

Similar to the effect of increasing the driving amplitude, the system response begins to oscillate at the driving frequency when ω_d is close to ω_n. This phenomenon is known as entrainment, or suppression (of the natural response). The range of frequencies for which this is true, known as the basin of entrainment or region of suppression (and is bounded by the pull-out frequencies), varies with other factors such as driving amplitude.

Simulations of Limit Cycle Suppression

In order to investigate the steady-state suppression characteristics of the Van der Pol Oscillator, a set of calculations was performed to determine the strength of the driven frequency component, ω_{d}, and that of the natural frequency component, ω_{n}, as a function of the input amplitude of the excitation at the driving frequency. This is plotted for two values of the driving frequency relative to the natural frequency in Figure E.1. These clearly show suppression of the natural frequency response as the excitation level is increased, and compression of the driven response at higher levels. Note that, within the range and resolution of these simulations, the frequency of the limit cycle is unaffected by such suppression.

Figure E.1 shows a widening of the basin of entrainment as driving A is increased. The boundary between moderate and high driving levels is termed the pull-out amplitude; the value of this quantity is dependent upon a number of factors. As the driving frequency becomes more distant from the natural frequency of the oscillator, stronger driving amplitudes will be required to achieve the same level of suppression.

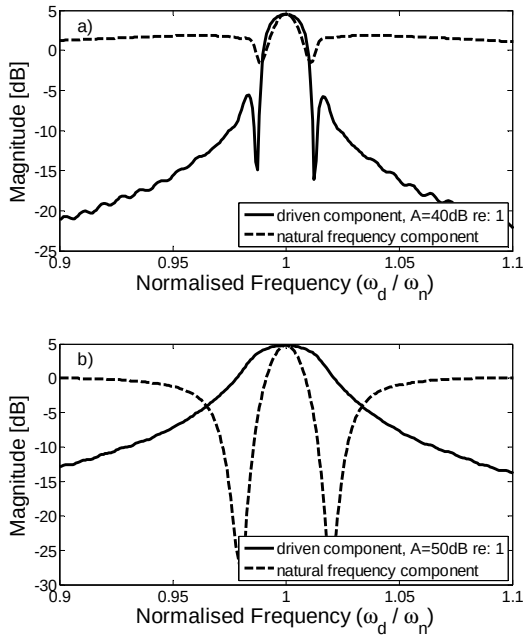


Figure E.1.a-b: Illustration of basin of entrainment. A solid line shows the driven frequency component, whereas a dashed line shows the natural frequency component.

E.2 Discussion

There are a number of interesting shared features in comparing the response of an active cochlea to that of a Van der Pol Oscillator. These include the ability to spontaneously oscillate and reach a limit cycle oscillation, and suppression and entrainment of this natural response given an external tonal source across a basin of attraction. Such phenomena have been illustrated in this report. However, there are also a number of features of cochlear response that cannot be accounted for by a Van der Pol Oscillator in isolation.

The features of cochlear activity that cannot be simulated by a Van der Pol Oscillator in isolation include: an asymmetrical basin of attraction at frequencies close to the oscillator's natural frequency; frequency 'pushing,' where the frequency of the SOAE moves away from the suppresser frequency; and a fluctuation in the limit cycle amplitude and frequency over time. The investigation of the basin of attraction of a Van der Pol Oscillator carried out in this report is largely restricted to frequencies within ±20% of the natural frequency. Outside this frequency range, the Van der Pol Oscillator appears to

possess asymmetrical properties. This is believed to be due to the nature of the energy-injecting mechanism that drives a Van der Pol Oscillator's limit cycle oscillations.

In contrast, it is not surprising that the basin of attraction in a live cochlea is not symmetrical. The mechanical properties of the basilar membrane vary with position down the cochlea; when the cochlear response is coupled together through the fluid, it is natural that the suppression would not be symmetrical about a given characteristic place. The variation of cochlear parameters may also be responsible for the observed phenomenon of downwards 'pushing' in frequency of an SOAE due to an external source.

List of References

Allen, J. B. (1980). 'Cochlear micromechanics—a physical model of transduction,' J. Acoust. Soc. Am. 68, 1660-1670.

Allen, J.B., and Fahey, P.F. (1992). 'Using acoustic distortion products to measure the cochlear amplifier gain on the basilar membrane,' J. Acoust. Soc. Am., 92: 178-188.

Ashmore, J.F. (1987). 'A fast motile response in guinea-pig outer hair cells: the cellular basis of the cochlear amplifier.' J. Physiol.; 388: 323-347.

Békésy, G. von. (1949). 'The vibration of the cochlear partition in anatomical preparation and in models of the inner ear.' J. Acoust. Soc. Am., 21, 233-245.

Bell, A. (2001). 'Helmholtz's piano strings: reverberation of ripples on the tectorial membrane,' Retrieved 22nd September, 2008, from: http://cogprints.org/1735/0/Helmholtz.pdf

de Boer, E ., and Viergever, M. (1982). 'Validity of the Lieuville-Green (or WKB) method for cochlea mechanics,' Hear. Res. 8, 131 - 155.

de Boer, E. (1991). 'Auditory Physics. Physical principles in hearing theory. III.' Phys. Rep. 203:127-229.

de Boer, E. (1995). 'The 'inverse problem' solved for a three-dimensional model of the cochlea. I. Analysis.' J. Acoust. Soc. Am. 98: 896-903.

List of References

de Boer, E. (1996). 'Mechanics of the Cochlea: Modeling Efforts.' In: The cochlea, edited by P. Dallos, A. N. Popper, and R. R. Fay. Springer-Verlag, New York, 258-317.

Brass, D., and Kemp, D.T., (1993). 'Suppression of stimulus frequency otoacoustic emissions.' J. Acoust. Soc. Am. 93: 920–939

Braun, M. (1997). 'Frequency spacing of multiple spontaneous otoacoustic emissions shows relation to critical bands: a large-scale cumulative study,' Hear. Res., 114, 197-203.

Brownell, W.E., Bader, C.R., Bertrand, D., and Ribaupierre, Y. de (1985). 'Evoked mechanical responses of isolated cochlear outer hair cells.' Science 227: 194-196.

Burns, E.M., Strickland, E.A., Tubis, A. and Jones, K. (1984). 'Interactions among spontaneous otoacoustic emissions. I. Distortion products and linked emissions.' Hear. Res. 16: 271-278.

Cody, A.R., and Russel, I.J. (1987). 'The responses of hair cells in the basal turn of the guinea pig cochlea to tones.' J. Physiol. 383, 551-569.

Coleman, J. (2008). '2. Speaking and Hearing.' Retrieved 22nd September, 2008, from: http://www.phon.ox.ac.uk/~jcoleman/speaking_hearing.htm

Cooper, N.P. (1998). 'Harmonic distortion on the basilar membrane in the basal turn of the guinea-pig cochlea,' J. Physio. 509: 277-288.

Dallmayr, C. (1985). 'Spontane oto-akusticsche Emissionen: Statistik und Reaktion auf akustische Störtöne.' Acustica 59, 67-75.

Dallmayr, C. (1986). 'Stationäre und dynamische Eigenschaften spontaner und simultan evozierter oto-akustischer Emissionen. Dissertation, Technische Universität, Munich.

Diependaal, R.J., de Boer, E., Viergever, M.A. (1987). 'Cochlear power flux as an indicator of mechanical activity,' J. Acoust. Soc. Am. 82 (3), 917-926.

van Dijk, P. and Wit, H.P., (1990). 'Synchronization of spontaneous otoacoustic emissions to a 2f1-f2 distortion product.' J. Acoust. Soc. Am. 88: 850–85

Duifhuis, H., Hoogstraten, H.W., van Netten, S.M., Diependaal, R.J. and Bialek, W. (1985) 'The cochlear partition with coupled van der Pol oscillators.' In: Peripheral Auditory Mechanisms, Edited by: J.B. Allen, J.L. Hall, A. Hubbard, S.T. Neely and A. Tubis. Berlin: Springer, 290-295.

Duke, T., and Jülicher, F. (2003). 'The active travelling wave in the cochlea.' Phys. Rev. Lett. 90.

Elliott, S.J., Ku, E.M., and Lineton, B. (2007). 'A state space model for cochlear mechanics,' J. Acoust. Soc. Am. 122(5), 2759-2771.

Fletcher, N.H., and Rossing, T.D. (1998). 'the Physics of Musical Instruments,' New York: Springer-Verlag.

Franklin, G.F., Powell, J.D., and Emami-Naeini, A. (1991). Feedback Control of Dynamic Systems. U.S.A.: Addison-Wesley.

Fukazawa, T. (2002). 'How can the cochlear amplifier be realized by the outer hair cells which have nothing to push against?' Hear Res., 172: 53-61.

Fukazawa, T. and Tanaka, Y. (1996). 'Spontaneous otoacoustic emissions in an active feed-forward model of the cochlea.' Hear. Res. 95: 135-143.

Gelfand, S.A. (1998). Hearing: an introduction to psychological and physiological acoustics. New York: Marcel Dekker.

Geisler, C.D., and Sang, C. (1995). 'A cochlear model using feed-forward outer-hair-cell forces.' Hear. Res., 86: 132-146.

Gold, T. (1948). 'Hearing. II. The physical basis of the action of the cochlea.' Proc. R. Soc. Lond. B. Biol. Sci., 135:492–498.

List of References

Greenwood, D. D. (1990). 'A cochlear frequency-position function for several species---29 years later,' J. Acoust. Soc. Am. 87: 2592.

Guinan, J.J., Jr. Backus, B.C., Lilaonitkul, W., and Aharonson, V. (2003). 'Medial olivocochlear efferent reflex in humans: otoacoustic emission (OAE) measurement issues and the advantages of stimulus frequency OAEs.' J. Assoc. Res. Otolaryngol., 4: 521-40.

Guinan, J.J., Lin, T. and Cheng, H. (2005). 'Medial-olivocochlear-efferent inhibition of the first peak of auditory-nerve responses: Evidence for a new motion within the cochlea.' J. Acoust. Soc. Am. 118: 2421-2433.

Hall, J.W. (2000). Handbook of Otoacoustic Emissions. San Diego: Singular Publishing Group.

Hardt, D.E., (2008). Understanding Poles and Zeros. Retrieved 22nd September, 2008, from: http://www.mit.edu/2.14/www/Handouts/PoleZero.pdf

Harte, J.M. (2004). Nonlinear system identification for otoacoustic emissions. Doctoral thesis. Southampton, U.K.: Southampton University.

Helmoltz, H.F. von (1874). Théorie Physiologique de la Musique, Masson G. Ed., Paris.

van Hengel, P.W.J., Duifhuis H., 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(6), 3566-3571.

Holley, M.C. (1996). 'Outer Hair Cell Motility,' in The cochlea, edited by P. Dallos, A. N. Popper, and R. R. Fay. Springer-Verlag, New York, 258-317.

How, J., and Lutman, M.E. (2007). 'Transient evoked otoacoustic emission input-output function variation in a large sample of neonates and implications for hearing screening.' Int. J. Aud., 46:670-679.

Hubbard, A.E. and Mountain, D.C. (1996). Models of the Cochlea. In: Auditory Computation, edited by Hawkins, H.L., McMullen, T.A., Popper, A.N., and Fay, R.R.. New York: Springer-Verlag, p. 62-120.

Jesteadt, W. (Editor) (1997). Modeling Sensorineural Hearing Loss. Hillsdale, NJ: Erlbaum.

Johnston, B.M., Patuzzi, R., and Yates, G.K. (1986). 'Basilar membrane measurements and the travelling wave.' Hear. Res., 22, 147-153.

Kalluri, R., and Shera, C.A. (2007). 'Near equivalence of human click-evoked and stimulus-frequency otoacoustic emissions.' J. Acoust. Soc. Am. 121: 2097-2110.

Kanis, L.J. and de Boer, E. (1993). 'Self-suppression in a locally active nonlinear model of the cochlea: A quasi-linear approach' J. Acoust. Soc. Am. 94: 3199-3206.

Kemp, D. T. (1979). ''Evidence of mechanical nonlinearity and frequency selective wave amplification in the cochlea,'' Arch. Oto-Rhino-Laryngol. 224, 37–45.

Kemp, D. T. (1986). "Otoacoustic emissions, travelling waves and cochlear mechanisms," Hear. Res. 22, 95–104.

Kemp, D. and Brill, O.J. (2008). '5 mHz oscillations in OAE intensity following sound exposure (A).' J. Acoust. Soc. Am., 123: 3852.

Kemp, D.T., and Chum, R.A. (1980). 'Properties of the generator of stimulated acoustic emissions.' Hear. Res., 2, 213-232.

Kemp DT, Ryan S, Brah P. (1990) A guide to the effective use of otoacoustic emissions. Ear Hear; 11: 93-105.

Kim, D.O., Neely, S.T., Molnar, C.E., Matthews, J.W. (1980). An active cochlear model with negative damping in the partition: comparison with Rhode's ante- and post-mortem observations. In: van den Brink G., Bilsen, F.A. (eds.) Psychophysical,

Physiological, and Behavioural Studies in Hearing. Delft: Delft University Press, 7-14.

Kinsler, L. E. (1982). Fundamentals of acoustics. New York: Wiley.

de Kleine, H.P. Wit, P. van Dijk and P. Avan, (2000). 'The behavior of spontaneous otoacoustic emissions during and after postural changes.' J. Acoust. Soc. Am. 107: 3308–3316.

Kolston, P.J. (1999). 'Comparing in vitro, in situ, and in vivo experimental data in a three-dimensional model of the mammalian cochlear mechanics.' Proc. Natl. Acad. Sci. U.S.A. 96, 2676-3681.

Kolston, P.J. (2000). 'The importance of phase data and model dimensionality to cochlear mechanics.' Hear. Res., 145: 25-36.

Kolston, P.J., Viergever, M.A., de Boer, E., Diependaal, R.J., (1989). 'Realistic mechanical tuning in a micromechanical cochlear model.' J. Acoust. Soc. Am. 86, 133-140.

Konrad-Martin, D. and Keefe, D.H. (2003). 'Time–frequency analyses of transient-evoked stimulus frequency and distortion-product otoacoustic emissions: testing cochlear model predictions,' J. Acoust. Soc. Am. 114: 2021–2043.

Konrad-Martin, D., and Keefe, D.H. (2005). 'Transient-evoked stimulus-frequency and distortion-product otoacoustic emissions in normal and impaired ears,' J. Acoust. Soc. Am. 117: 3799–3815.

Kringlebotn, M. (1988) 'Network Model for the Human Middle Ear.' Scand. Audiol 17: 75-85.

Kros, C.J., Rüsch, A. and Richardson, G.P. (1992). Mechanoelectrical transducer currents in hair cells of the cultured mouse cochlea. Proc. Royal Soc. 249:185-193.

List of References

Kros, C.J. (1996). Physiology of Mammalian Cochlear Hair Cells. In: The Cochlea, edited by Dallos, P., Popper, A.N., and Fay, R.R. New York: Springer-Verlag, p. 318-385.

Kruglov, A.V., Artamasov, S.V., Frolenkov, G.I., and Tavartkiladze, G.A. (1997). 'Transient evoked otoacoustic emission with unexpectedly short latency,' Acta Oto-Laryngologica, 117: 174-178.

Ku, E.M., Elliott, S.J. and Lineton, B. (2008). 'Statistics of instabilities in a state space model of the human cochlea,' J. Acoust. Soc. Am.; 124: 1068-1079.

Levine, W.S. (1996). The Control Handbook. New York: CRC Press.

Lim, D.J. (1986). 'Functional structure of the organ of Corti: a review.' Hear. Res., 22: 117-146.

Lineton, B. (2001). 'Testing a model of the stimulus frequency otoacoustic emissions in humans,' Ph.D. thesis, University of Southampton, Southampton, U.K.

Lonsbury-Martin, B.L., Martin, G.K., Probst, R., and Coats, A.C. (1987). 'Spontaneous otoacoustic emissions in a nonhuman primate. II. Cochlear anatomy,' Hear. Res., 33, 69-64.

Mammano, F., and Nobili, R. (1993). Biophysics of the cochlea: linear approximation. J Acoust Soc Am; 93: 3320-3332.

Martin,P., and Hudspeth, A.J. (1999). 'Active hair-bundle movements can amplify a hair cell's response to oscillatory mechanical stimuli,' Proc. Natl. Acad. Sci. U.S.A. 96, 14380-14385.

Mathworks. (2008). 1106 – Memory Management Guide. Retrieved 20th September, 2008, from Mathworks website: http://www.mathworks.com/support/tech-notes/1100/1106.html

Matthews, J. W., Ruggero, M. A., and Steele, C. R., editors, The Mechanics and Biophysics of Hearing, pages 324–332. Springer-Verlag, New York.

McFadden, D., and Mishra, R. (1993). 'On the relation between hearing sensitivity and otoacoustic emissions,' Hear. Res., 71, 208-213.

Meaud, J., and Grosh, K. (2008). 'Exploring the role of hair bundle motility using a macroscopic cohlear model,' Proc. 10th Int. Wkshp. on the Mech. of Hear., 201-202.

Moleti, A., Sisto, R., Tognola, G., Parazzini, M., Ravazzani, P., and Grandori, F. (2005). 'Otoacoustic emission latency, cochlear tuning, and hearing functionality in neonates,' J. Acoust. Soc. Am. 118: 1576–1584.

Murphy, W.J., Tubis, A., Talmadge, C.L. and Long, G.R.. (1995). 'Relaxation dynamics of spontaneous otoacoustic emissions perturbed by external tones. II. Suppression of interacting emissions,' J. Acoust. Soc. Am. 97: 3711–3720.

Neely, S.T. (1981). 'Finite difference solution of a two-dimensional mathematical model of the cochlea,' J. Acoust. Soc. Am. 69(5), 1386-1393.

Neely, S.T., and Allen, J.B. (2008). 'Retrograde waves in the cochlea,' Proc. 10th Int. Wkshp. on the Mech. of Hear., 257-262.

Neely, S.T., and Kim, D.O. (1983). 'An active cochlear model showing sharp tuning and high sensitivity,' Hear. Res. 9: 123–130.

Neely, S.T., and Kim, D.O. (1986). 'A model for active elements in cochlear biomechanics,' J. Acoust. Soc. Am. 79 (5), 1472-1480.

Neely, S.T., and Kim, D.O. (2007). 'Cochlear models incorporating active processes,' In: Active Processes and Otoacoustic Emissions in Hearing, edited by Manly, G.A., Fay, R.R., and Popper, A.N. New York: Springer-Verlag.

van Netten, S.M., and Duifhuis, H. (1983) 'Modelling an active, nonlinear cochlea', in Mechanics of Hearing, edited by E. de Boer and M.A. Viergever. Netherlands: Nijhoff/Delft Univ. Press, 143-151.

Newburg, S.O. and Mountain, D.C. (2008). 'Mechanical response of the basilar membrane to lateral micromanipulation,' Proc. 10th Int. Wkshp. on the Mech. of Hear. 136-141.

Nobili, R. and Mammano, F. (1996). 'Biophysics of the cochlea. II: Stationary nonlinear phenomenology.' J.Acoust.Soc.Am. 99: 2244-2255.

Nobili, R., Vetesnik, A., Turicchia, L., Mammano, F. (2003). 'Otoacoustic emissions from residual oscillations of the cochlear basilar membrane in a human ear model,' J. Assoc. Res. Otol., 4, 478-494.

Nuttall, A.L., Grosh, K., Zheng, J., de Boer, E., Zou, Y., Ren, T. (2004). 'Spontaneous basilar membrane oscillation and otoacoustic emission at 15kHz in a guinea pig,' J. Assoc. Res. Otol., 5, 337-348.

O'Conner, K.N. and Puria, S. (2008). 'Middle-ear circuit model parameters based on a population of human ears,' J. Acoust. Soc. Am., 123: 197-211.

Parthasarathi, A.A., Grosh, K., Nuttall, A.L. (2000). 'Three-dimensional numerical modelling for global cochlear dynamics.' J. Acoust. Soc. Am. 107: 474-485.

Pascal, J., A. Bourgeade, M. Lagier, and C. Legros. (1998). 'Linear and Nonlinear Model of the Human Middle Ear.' J. Acoust. Soc. Am. 104: 1509-1516.

Patuzzi, R. Cochlear Micromechanics and Macromechanics. In: The Cochlea, edited by Dallos, P., Popper, A.N., and Fay, R.R. New York: Springer-Verlag, 1996, p. 186-257.

Pélanová, J., Groh, D., Popear, J., Kabelka, Z., Syka, J. (2007). 'Presence and characteristics of spontaneous otoacoustic emissions in children and adolescents.' Poster presentation at Inner Ear Biology 2007.

Penner, M.J., and Zhang, T. (1997). 'Prevelance of spontaneous otoacoustic emissions in adults revisted,' Hear. Res., 103, 28-34.

Phillips, C.L., and Harbor R.D. (2000). Feedback control systems. 4th ed. Upper Saddle River, N.J.: Prentice Hall.

Pickles, J.O. (2003). 'An introduction to the physiology of hearing.' Second Edition. London: Academic Press.

Probst, R., Coots, A. C., Martin, G. K., and Lonsbury-Martin, B. L. (1986). 'Spontaneous, click-, and toneburst-evoked otocoustice missions from normal ears,' Hear. Res. 21: 261-275.

Probst, R., Lonsbury-Martin, B.L., and Martin, G.K. (1991). 'A review of otoacoustic emissions,' J. Acoust. Soc. Am. 89(5), 2027-2066.

Puria, S. (2003). 'Measurements of human middle ear forward and reverse acoustics: Implications for otoacoustic emissions,' J. Acoust. Soc. Am. 113, 2773-2789.

Recio, A., and Rhode, W.S. (2000). 'Basilar membrane responses to broadband stimuli.' J. Acoust. Soc. Am., 108: 2281-2298.

Recio, A., Rich, N.C., Narayan, S.S., and Ruggero, M. (1998). 'Basilar-membrane responses to clicks at the base of the chinchilla cochlea.' J. Acoust. Soc. Am. 103: 1972-1989.

Ren, T. (2002). 'Longitudinal pattern of basilar membrane vibration in the sensitive cochlea,' Proc. Nat. Acad. Sc., 99:17101-17106.

Rhode, W.S. (1971). 'Observations of the vibration of the basilar membrane in squirrel monkeys using the Mössbauer technique.' J. Acoust. Soc. Am., 49:1218-1231.

Rhode, W.S., Robles, L. (1974). 'Evidence for nonlinear vibration in the cochlea from Mössbauer investigations.' J. Acoust. Soc. Am., 55:588-596.

Rhode, W.S. (1984). 'Cochlear mechanics.' Ann. Rev. Physiol. 46:231-246.

Robles, L., and Ruggero, M.A. (2001). 'Mechanics of the mammalian cochlea,' Phys. Rev., 81(3), 1305-1352.

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. (2001a). 'Frequency glides in the click-response of the basilar membrane and auditory nerve: Their scaling behaviour and origin in traveling-wave dispersion,' J. Acoust. Soc. Am. 109, 2023-2034.

Shera, C.A. (2001b). 'Intensity-invariance of fine time structure in basilar-membrane click-responses: Implications for cochlear mechanics,' J. Acoust. Soc. Am. 110 332-348.

Shera, C.A. (2003). 'Mammalian spontaneous otoacoustic emissions are amplitude-stabilized cochlear standing waves,' J. Acoust. Soc. Am. 114, 244-262.

Shera, C.A. (2007). 'Laser amplification with a twist: Traveling-wave propagation and gain functions throughout the cochlea,' J. Acoust. Soc. Am. 122, 2738-2758.

Shera, C.A., and Guinan, J.J. (1999). 'Evoked otoacoustic emissions arise by two fundamentally different mechanisms: A taxonomy for mammalian OAEs,' J. Acoust. Soc. Am. 105, 782.

Shera, C.A., and Guinan, J.J. (2003). 'Stimulus-frequency-emission group delay: A test of coherent reflection filtering and a window on cochlear tuning.' J. Acoust. Soc. Am., 113: 2762–2772

Shera, C.A., Tubis, A., and Talmadge, C.L., (2003). 'Stimulus-spectrum irregularity and the generation of evoked and spontaneous otoacoustic emissions: comments on the model of Nobili et al.' Last viewed 17/09/2008. http://otoemissions.org/whitepapers/biophysics/chris_shera.html

Shera, C.A., Tubis, A., and Talmadge, C.L. (2005). 'Coherent reflection in a two-dimensional cochlea: Short-wave versus long-wave scattering in the generation of reflection-source otoacoustic emissions.' J. Acoust. Soc. Am., 118: 287–313.

Shera, C.A., and Zweig, G. (1991a). 'Reflection of retrograde waves within the cochlea and at the stapes.' J. Acoust. Soc. Am., 89:1290–1305.

Shera, C.A., and Zweig, G. (1991b). 'A symmetry suppresses the cochlear catastrophe.' J. Acoust. Soc. Am., 89:1276–1289

Shera, C. A., and Zweig, G. (1993). 'Order from chaos: resolving the paradox of periodicity in evoked otoacoustic emission,' Biophysics of Hair Cell Sensory Systems, edited by H. Duifhuis, J. W. Horst, P. van Dijk, and S. M. van Netten (World Scientific), Singapore, pp. 54–63.

Sisto, R., and Moleti, A. (2002). 'On the frequency dependence of the otoacoustic emission latency in hypoacoustic and normal ears,' J. Acoust. Soc. Am. 111, 297–308.

Sisto, R., and Moleti, A. (2007). 'Transient evoked otoacoustic emission latency and cochlear tuning at different stimulus levels,' J. Acoust. Soc. Am. 122, 2183-2190.

Sisto, R., Moleti, A., and Lucertini, M. (2001). 'Spontaneous otoacoustic emissions and relaxation dynamics of long decay time OAEs in audiometrically normal and impaired subjects.' J. Acoust. Soc. Am., 109: 638-647.

Sisto, R., Moleti, A. and Shera, C.A. (2007). 'Cochlear reflectivity in transmission-line models and otoacoustic emission characteristic time delays.' J. Acoust. Soc. Am., 122: 3554-3561.

Steele, C.R., Baker, G., Tolomeo, J., Zetes, D. (1993). 'Electro-mechanical models of the outer hair cell.' In Duifhuis H, Horst JW, van Dijk P, van Netten SM (eds): Biophysics of Hair Cell Sensory Systems. World Scientific, 207-215.

Strube, H.W. (1989). 'Evoked otoacoustic emissions as cochlear Bragg reflections,' Hear. Res., 38, 35-46.

Talmadge, C.L., Long, G.R., Murphy, W.J., Tubis, A. (1993). 'New off-line method for detecting spontaneous otoacoustic emissions in human subjects,' Hear. Res., 71, 170-182.

Talmadge, C. L. and A. Tubis (1993). 'On modeling the connection between spontaneous and evoked otoacoustic emissions.' In H. Duifhuis, J. W. Horst, P. van Dijk, and S. M. van Netten (Eds.), Biophysics of Hair Cell Sensory Systems, pp. 25–32. Singapore: World Scientific.

Talmadge, C.L. and Tubis, A., Long, G.R., and Piskorski, P. (1998). 'Modeling otoacoustic emission and hearing threshold fine structures,' J. Acoust. Soc. Am. 104: 1517-1543.

Talmadge, C.L. and 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 structure,' J. Acoust. Soc. Am. 108: 2911-2932.

Tognola G, Grandori F, Avan P, Ravazzani P, Bonfils P. (1999). 'Frequency-specific information from click evoked otoacoustic emissions in noise-induced hearing loss.' Audiology; 38: 243-50.

Whitehead, M.L. (1988). 'The interactions and instability of spontaneous otoacoustic emissions (SOAEs): clues to the nature of the emission generator.' Brit. J. Audio. 23, 172.

Wit, H.P. (1986). 'Statistical properties of a strong spontaneous otoacoustic emission,' in Peripheral Auditory Mechanisms, edited by J.B. Allen, J.L. Hall, A.E. Hubbard, S.T. Neely, and A. Tubis (Springer-Verlag, Berlin), 221-228.

Wit, H.P, van Dijk, P., and Avan, P. (1994). 'Wavelet analysis of real and synthesised click evoked otoacoustic emissions,' Hearing Res., 73, 141-147.

Xin, J., Qi, Y., and Deng, L. (2003). 'Time domain computation of a nonlinear nonlocal cochlear model with applications to multitone interaction in hearing,' Comm. Math. Sci., 2: 211-227.

Zweig, G. (1991). 'Finding the impedance of the organ of Corti,' J. Acoust. Soc. Am. 89, 1229–1254.

Zweig, G., and Shera, C.A. (1995). 'The origin of periodicity in the spectrum of evoked otoacoustic emissions,' J. Acoust. Soc. Am. 98, 2018-2047.

Zwicker, E. (1990). 'On the frequency separation of simultaneously evoked otoacoustic emissions' consecutive extrema and its relation to cochlear travelling waves,' J. Acoust. Soc. Am., 88.

Zwicker, E., and Peisl W. (1990). 'Cochlear preprocessing in analog models, in digital models, and in human inner ear,' Hear. Res. 44, 209-216.

Zwicker, E., and Schloth, E. (1984). 'Interrelation of different oto-acoustic emissions.' J. Acoust. Soc. Am. 75(4), 1148-1154.

Zwislocki, J. 'Analysis of the Middle-Ear Function. Part I: Input Impedance.' J Acoust Soc Am 34 (1962): 1514-1523.

Zwislocki, J.J., and Kletsky, E.J. (1979). 'Tectorial membrane: a possible effect on frequency analysis in the cochlea.' Science, 204:639-641.