

# Spatiotemporal coding of sound in the auditory nerve for cochlear implants

Ian Christopher Bruce  
BE (Melb)

Submitted in total fulfilment of the requirements  
of the degree of Doctor of Philosophy

December 1997

Department of Otolaryngology  
The University of Melbourne

# Abstract

In this thesis spatiotemporal patterns of auditory nerve activity are modelled, in order to determine how they may code features of acoustical and electrical stimuli. In particular, questions and applications of interest to cochlear implant development are investigated, including temporal coding of stimulus features in populations of nerve fibres and stochastic activity in responses to electrical stimulation.

A model of acoustical stimulation of the auditory nerve is developed to investigate how interspike intervals across a pair of auditory nerve fibres may code acoustical stimulus frequency. The results indicate that such frequency coding may take place across fibres for frequencies below 1–3 kHz, if the fibres originate from nearby sites in the cochlea. This investigation is then extended to spatiotemporal combination, specifically spatiotemporal summation, of activity in more than two fibres. A model is developed to determine theoretically how various degrees of auditory nerve fibre convergence onto summing neurons in the cochlear nucleus would affect estimation and discrimination of stimulus features from the summing neurons’ outputs. Results obtained from the model indicate that for coding of the stimulus frequency in the periodicity of firings, the optimal spread of input fibres is very narrow and temporal integration is very short. In contrast, cochlear nucleus cell types which have a wide spread of inputs and a very short temporal integration window may be useful in performing estimation and discrimination tasks based on the average discharge rate. The first two investigations are of responses to pure-tone stimuli in identical auditory nerve fibres—the possibility of coding different complex-stimulus features in fibres with different response properties is therefore investigated next. A model of acoustical stimulation is developed to determine an expression for the Cramer–Rao bound for frequency estimation of the envelope and fine structure of complex sounds by groups of fibres with different response properties. The estimation variances are calculated for some typical estimation tasks and demonstrate how, in the examples studied, a combination of low and high threshold fibres may improve the estimation performance of an ‘efficient’ observer.

A stochastic model of auditory response to single electrical pulses is also derived, and direct comparisons are made between this model and the same model without the stochastic component. In contrast to the deterministic model, the stochastic model accurately predicts probabilities of

discharge measured in response to single biphasic pulses. A model of loudness in cochlear implant users is developed, which includes spatiotemporal summation of auditory nerve activity and can incorporate either the deterministic or the stochastic description of auditory nerve response. For all parameters investigated, the inclusion of stochastic activity in the model is found to produce more accurate predictions of behavioural performance than it is not included. The stochastic model of auditory nerve response to electrical stimulation is then extended to describe responses to pulse-train stimuli. The stochastic model, unlike the deterministic model, accurately predicts means and variances of discharge rates measured in response to pulse trains. The pulse-train stochastic model provides a means for extending the loudness model to predict psychophysical data for higher pulse rates.

This is to certify that:

- (i) the thesis comprises only my original work;
- (ii) due acknowledgement has been made in the text to all other material used; and
- (iii) the thesis is less than 100 000 words in length, exclusive of tables, maps, bibliographies, appendices and footnotes.

# Acknowledgements

I wish to thank my thesis supervisors, Dr. Laurence Irlicht and Prof. Graeme Clark, for their support, guidance, encouragement and example over the past three and a half years. I am indebted to Laurence for his unwavering commitment to me and my research efforts, for all he has taught me about the scientific process and doing it well, for his thorough and patient supervision and editing of all I have written and presented, and for his care and friendship. I would like to thank Graeme for his support of mathematical modelling in cochlear implant research, for providing a stimulating research environment, for the example he gives in his commitment to assisting the hearing impaired, and for all he has taught me about communicating well to a multi-disciplinary audience.

Much thanks also goes to Dr. Mark White for his collaboration and friendship during his year as Visiting Professor in our department and since. I am extremely grateful for his encouragement of my efforts in our collaboration, for his passion about science and his humbleness in passing on what he knows, and for his open friendliness.

I would like to thank the other co-authors who have contributed to the presentations and publications which have come out of this research: Mr. David Au, Dr. Stephen O’Leary, Dr. Scott Dynes and Dr. Eric Javel; and those who have given comments on the preparation this thesis: Dr. Anthony Burkitt and Prof. Barbara Cone-Wesson. My thanks is also extended to all in the Department of Otolaryngology and associated research groups who have befriended, taught, and helped me in my time as a student.

I am beholden to my friends and family for their love and care. Foremost, I wish to thank my wife Gillian, who has been a continual source of support and encouragement in our two years of marriage, and has helped keep my life in perspective by her great love and friendship.

For all of the above and more I thank God—for the life that he has given me in Jesus and for the many blessings I’ve received during my Ph.D. candidature.

This work was funded by the Human Communication Research Centre, the Department of Otolaryngology, The University of Melbourne, and The Bionic Ear Institute.

# Contents

<b>Abstract</b>	<b>i</b>
<b>Acknowledgements</b>	<b>iv</b>
<b>List of Figures</b>	<b>viii</b>
<b>Preface</b>	<b>xvi</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Anatomy and physiology of the ear . . . . .	2
1.2 Coding of sound in the normal ear . . . . .	7
1.2.1 Spatial coding . . . . .	9
1.2.2 Temporal coding . . . . .	9
1.2.3 Spatiotemporal coding . . . . .	10
1.3 Cochlear implants . . . . .	10
1.4 Thesis overview . . . . .	14
<b>2 Frequency coding by cross-fibre interspike intervals</b>	<b>16</b>
2.1 Introduction . . . . .	16
2.2 The cochlear model . . . . .	17
2.3 The hair-cell/auditory-nerve model . . . . .	20
2.4 Cross-fibre interspike interval probability distribution . . . . .	21
2.5 Results . . . . .	22
2.6 Discussion . . . . .	28
<b>3 Spatiotemporal summation of auditory nerve firings</b>	<b>29</b>
3.1 Introduction . . . . .	29
3.2 Statistical model of summed AN response . . . . .	30
3.2.1 Single AN fibre model . . . . .	31
3.2.2 Fitting the single AN fibre model to experimental data . . . . .	33

3.2.3	Calculation of the variance in single fibre response . . . . .	38
3.2.4	A summing-neuron model . . . . .	40
3.3	Response properties of the summing-neuron model . . . . .	43
3.3.1	Synchronisation in the model output . . . . .	43
3.3.2	Mean-to-standard-deviation ratio . . . . .	47
3.4	Application of the model to theories of sound coding . . . . .	49
3.4.1	Temporal (time-period) frequency coding . . . . .	52
3.4.2	Average-rate intensity coding . . . . .	52
3.5	Discussion . . . . .	55
<b>4</b>	<b>The effect of neural thresholds on frequency estimation</b>	<b>56</b>
4.1	Introduction . . . . .	56
4.2	Signal and network models . . . . .	57
4.2.1	Signal model . . . . .	57
4.2.2	Neural network model . . . . .	57
4.2.3	Filtering properties of the neural model . . . . .	59
4.2.4	Cramer–Rao bounds for neural estimation of frequency . . . . .	60
4.2.5	Interspike interval analysis . . . . .	64
4.3	Results . . . . .	64
4.3.1	Estimation by an efficient estimator . . . . .	65
4.3.2	Estimation from the interspike-interval distribution . . . . .	66
4.4	Discussion . . . . .	67
<b>5</b>	<b>A stochastic model of the electrically stimulated auditory nerve: Single-pulse response</b>	<b>68</b>
5.1	Introduction . . . . .	68
5.2	Methods . . . . .	72
5.2.1	Physiological data . . . . .	72
5.2.2	Model of AN response . . . . .	73
5.3	Single-pulse response . . . . .	73
5.3.1	Deterministic model of single-pulse response . . . . .	74
5.3.2	Prediction of single-pulse response by deterministic model . . . . .	76
5.3.3	Stochastic model of single-pulse response . . . . .	77
5.3.4	Prediction of single-pulse response by stochastic model . . . . .	79
5.4	Total auditory nerve response . . . . .	80
5.4.1	Fitting of model parameters to AN statistics from cat . . . . .	80
5.4.2	Current spread . . . . .	87

5.4.3	Output of total auditory nerve model . . . . .	87
5.5	Discussion . . . . .	89
<b>6</b>	<b>The effects of stochastic neural activity in a model predicting loudness perception with cochlear implants</b>	<b>95</b>
6.1	Introduction . . . . .	95
6.2	Methods . . . . .	97
6.2.1	Psychophysical model . . . . .	97
6.2.2	Determination of psychophysical measures . . . . .	99
6.2.3	Psychophysical data . . . . .	104
6.3	Results . . . . .	105
6.3.1	Model response properties . . . . .	105
6.3.2	Threshold versus phase duration (pulse width): Strength-duration . . . . .	108
6.3.3	Threshold and uncomfortable loudness versus phase duration: Dynamic range	108
6.3.4	Effect of electrode configuration on dynamic range . . . . .	111
6.3.5	Effect of electrode configuration on strength-duration . . . . .	111
6.3.6	Threshold versus number of pulses: Temporal integration . . . . .	113
6.3.7	Intensity Difference Limen (IDL) . . . . .	113
6.3.8	Effect of number of fibres on threshold, dynamic range and intensity difference limen . . . . .	117
6.4	Discussion . . . . .	118
6.4.1	Using the model to understand psychophysical behaviour . . . . .	118
6.4.2	Applicability of the model . . . . .	120
6.4.3	Model extensions . . . . .	121
<b>7</b>	<b>A stochastic model of the electrically stimulated auditory nerve: Pulse-train response</b>	<b>124</b>
7.1	Introduction . . . . .	124
7.2	Methods . . . . .	125
7.2.1	Physiological data . . . . .	125
7.2.2	Model of AN response . . . . .	125
7.3	Pulse-train response . . . . .	125
7.3.1	Stochastic model of pulse-train response . . . . .	126
7.3.2	Analytical approximation of stochastic model . . . . .	128
7.3.3	Analysis of the analytical approximation's accuracy . . . . .	134
7.3.4	Prediction of pulse-train response . . . . .	138
7.3.5	Variance of discharge rate . . . . .	144

7.4	Total auditory nerve response . . . . .	147
7.5	Discussion . . . . .	151
<b>8</b>	<b>Conclusions</b>	<b>154</b>
8.1	Summary . . . . .	154
8.2	Implications for sound coding by cochlear implants . . . . .	157
8.2.1	Spatiotemporal combination of auditory nerve activity . . . . .	157
8.2.2	Effect of neural thresholds on frequency estimation . . . . .	158
8.2.3	Stochastic auditory nerve response to electrical stimulation . . . . .	159
8.2.4	Future directions . . . . .	163
<b>A</b>	<b>Proofs for Chapter 3</b>	<b>165</b>
A.1	Proof of Claim 3.2.1 . . . . .	165
A.2	Proof of Lemma 3.2.2 . . . . .	165
A.3	Proof of Lemma 3.3.1 . . . . .	166
<b>B</b>	<b>Proofs for Chapter 5</b>	<b>167</b>
B.1	Proof of Lemma 5.3.1 . . . . .	167
<b>C</b>	<b>Proofs for Chapter 6</b>	<b>168</b>
C.1	Proof of Lemma 6.2.1 . . . . .	168
<b>D</b>	<b>Proofs for Chapter 7</b>	<b>169</b>
D.1	Proof of Lemma 7.3.1 . . . . .	169
D.2	Proof of Lemma 7.3.2 . . . . .	170
D.3	Proof of Lemma 7.3.3 . . . . .	170
D.4	Proof of Lemma 7.3.4 . . . . .	171
D.5	Proof of Lemma 7.3.5 . . . . .	174
	<b>Bibliography</b>	<b>175</b>

# List of Figures

1-1	Anatomy of the human ear: <b>(A)</b> pinna; <b>(B)</b> external auditory meatus; <b>(C)</b> tympanic membrane; <b>(D)</b> malleus; <b>(E)</b> incus; <b>(F)</b> stapes; <b>(G)</b> oval window; <b>(H)</b> cochlea; <b>(I)</b> auditory nerve. . . . .	3
1-2	Schematic of the vibrations in the cochlear duct, which is depicted as if it had been unrolled. From Pickles (1982). . . . .	4
1-3	A cross-section of the cochlea. From Pickles (1982). . . . .	5
1-4	A cross-section of the organ of Corti. From Pickles (1982). . . . .	6
1-5	Instantaneous probability of firing (indicated by the height of the trace) for 19 auditory nerve fibres versus origin on the basilar membrane (distance from the stapes) and time. The basilar membrane displacement at two different times (dashed and solid lines) is shown on the left, indicating the movement of the travelling wave which produces this neural response. . . . .	8
1-6	Components of a typical cochlear implant: <b>(a)</b> microphone; <b>(b)</b> behind-the-ear speech processor; <b>(c)</b> body-worn speech processor; <b>(d)</b> transmitter coil; <b>(e)</b> receiver/stimulator; <b>(f)</b> electrode array; <b>(g)</b> cochlea; <b>(h)</b> auditory nerve. . . . .	11
2-1	Neely and Kim's (1986) lumped component model of cochlear micromechanics. The mass $M_1$ represents a cross section of the organ of Corti which is attached to rigid bone by stiffness and damping components $K_1$ and $R_1$ . The basilar membrane is "driven" by fluid pressure difference $P_d$ and active pressure source $P_a$ . The mass $M_2$ represents a cross section of the tectorial membrane which is attached to rigid bone by $K_2$ and $R_2$ . The two masses are coupled by $K_3$ and $R_3$ . $\xi_t$ represents the displacement of the tectorial membrane ( $M_2$ ), and $g\xi_b$ represents the displacement of the basilar membrane ( $M_1$ ). . . . .	18
2-2	Magnitude (top) and phase (bottom) of the hair-bundle vibration versus normalised position on the basilar membrane (distance from stapes), for a 0 dB SPL pure tone at five different frequencies. . . . .	19

2-3	Cross-fibre interspike interval probability distribution with both fibres at the characteristic place (solid lines), and with one fibre at the characteristic place and a second at a more apical site (dashed lines), for a 500 Hz (top) and 10 kHz (bottom) pure-tone. . . . .	22
2-4	Cross-fibre Synchronisation Index (SI) versus characteristic frequency, for two fibres at the characteristic place. . . . .	24
2-5	Cross-fibre SI (where one fibre is at the CP) versus the characteristic frequency of the second fibre. . . . .	25
2-6	Cross-fibre SI (where one fibre is at the CP) versus the characteristic place of the second fibre (normalised distance from stapes). . . . .	26
2-7	Cross-fibre SI (where one fibre is at the CP) versus the phase difference between the two fibres. . . . .	27
3-1	Basilar membrane vibration amplitude (top) and phase lag (bottom) curves for five frequencies between 500 Hz and 10 kHz. . . . .	35
3-2	$A_n$ (solid lines) and $B_n$ (dashed lines) versus normalised position on the basilar membrane for five different stimulus frequencies, as labelled in pane (c), at sound levels of $-30$ dB (a) to $60$ dB (d). . . . .	37
3-3	Analytical results: (a) PSI and (b) $\text{PSI}/\text{SI}_{\text{CP}}$ versus summation distance and tone frequency, for a spatiotemporal summing neuron centred at the CP and with $\tau \rightarrow 0$ . . . . .	45
3-4	Computer simulation results: (a) PSI and (b) $\text{PSI}/\text{SI}_{\text{CP}}$ versus summation distance and tone frequency, for a spatiotemporal summing neuron centred at the CP and with $\tau \rightarrow 0$ . . . . .	46
3-5	Expected counts (solid lines) and confidence intervals of at least 75% (dotted lines) for the number of spikes counted in 5 ms by a summing neuron with 10 input fibres. Top: Zero PSI. Bottom: Non-zero PSI. . . . .	48
3-6	Detectability versus discharge rate and integration time. . . . .	50
3-7	Detectability versus single-fibre discharge rate and number of input fibres. . . . .	51
3-8	Zero-PSI widths (o) and typical critical bandwidths (*) from a review by Moore (1977, p. 103). Top: Widths versus CP; Bottom: Widths (expressed in terms of the corresponding difference in CF) versus CF. . . . .	54
4-1	Filter characteristics of the cochlea to a 700 Hz tone. . . . .	58
4-2	Left: Ex. 1—Slightly modulated $s(t)$ . Sigmoids with optimal thresholds for 600, 700 and 100 Hz (top to bottom). Right: Ex. 2—Highly modulated $s(t)$ . Sigmoids with optimal thresholds for 600, 700 and 100 Hz (top to bottom). . . . .	60

4-3	Ex. 1: Slightly modulated $s(t)$ . Output values of the sigmoids with optimal thresholds for 600, 700 and 100 Hz. . . . .	61
4-4	Ex. 2: Highly modulated $s(t)$ . Output values of the sigmoids with optimal thresholds for 600, 700 and 100 Hz. . . . .	62
4-5	Left: Ex. 1—Slightly modulated $s(t)$ . Interspike-interval distribution for low-threshold ( $T = 1.0$ ) and high-threshold ( $T = 1.8$ ) sigmoids (top and bottom). Right: Ex. 2—Highly modulated $s(t)$ . Interspike-interval distribution for low-threshold ( $T = 1.0$ ) and high-threshold ( $T = 1.8$ ) sigmoids (top and bottom). . . . .	66
5-1	Membrane-potential traces from Verveen and Derksen (1968) showing fluctuations in nerve fibre transmembrane potentials at the nodes of Ranvier. . . . .	70
5-2	Deterministic model of single-pulse response. . . . .	76
5-3	Deterministic model (solid line) fit to probability of discharge data (o) from Neuron 3-21 of Javel et al. (1987) for a single biphasic pulse of duration 100 $\mu$ s/phase ( <i>cf.</i> stochastic model fit to this data plotted in Figure 5-5). . . . .	77
5-4	Stochastic model of single-pulse response. . . . .	78
5-5	Stochastic model (solid line) fit to probability of discharge data (o) from Neuron 3-21 of Javel et al. (1987) for a single biphasic pulse of duration 100 $\mu$ s/phase ( <i>cf.</i> deterministic model fit to this data plotted in Figure 5-3). . . . .	79
5-6	Fits to fibre I/O functions from Javel et al. (1987) for a single pulse of duration 200 $\mu$ s/phase. . . . .	81
5-7	Fits to fibre I/O functions from Javel et al. (1987) for a single pulse of duration 400 $\mu$ s/phase. . . . .	82
5-8	Relative spread versus threshold for neurons from Javel et al. (1987) as labelled, in response to a single biphasic pulse of width 200 $\mu$ s/phase (x) or 400 $\mu$ s/phase (o). . . . .	83
5-9	Discharge probability versus level functions for single symmetric biphasic anodic/cathodic pulses of differing durations. For each duration, several measures of the probability of discharge were made at different stimulus levels. Individual measures consisted of 100 stimulus presentations and are represented by asterisks. Curves are error function fits determined by a least-squares method. All data are from a single unit. . . . .	84
5-10	Relative spread as a function of phase duration for three units from the same cat. The relative spread was determined as described in the text; individual measures are represented by symbols. . . . .	85
5-11	Fits to probability of discharge versus stimulus intensity data for pulse widths of 100 $\mu$ s/phase to 5 000 $\mu$ s/phase. . . . .	86
5-12	Attenuation of the stimulus across the cochlea for monopolar (solid line) and bipolar (dashed line) electrode configurations. . . . .	88

5-13	Number of discharges versus stimulus intensity from model of 10 000 fibres in response to a single biphasic pulse for monopolar (solid line) and bipolar (dashed line) stimulation. Error bars indicate $\pm 1$ std. . . . .	90
5-14	Distribution of total number of discharges predicted by 10 000 iterations of Monte Carlo simulations in response to a single biphasic pulse for two different intensities, with a Poisson distribution fit (o) for the case where the mean count less than 15 and a Gaussian (Normal) distribution fit (solid line) for the case where the mean count is greater than 15. . . . .	91
5-15	Figure 5-11 replotted with an extended ordinate. . . . .	93
6-1	Cochlear neural model and behavioural loudness model. . . . .	97
6-2	Psychometric functions for behavioural threshold predicted by the deterministic (dotted line) and stochastic (solid line) auditory nerve models. . . . .	102
6-3	Mean thresholds from 4 iterations of Monte Carlo simulations (o), with error bars indicating $\pm 1$ standard deviation, compared with thresholds obtained via the analytical method (lines). . . . .	103
6-4	Neural response versus stimulus intensity from model of 10 000 fibres in response to a single pulse of duration 100 and 2 000 $\mu\text{s}$ /phase (solid lines and dashed lines respectively). . . . .	106
6-5	Neural response versus stimulus intensity from model of 10 000 fibres in response to a single pulse for monopolar (solid line) and bipolar (dashed line) stimulation. . . . .	107
6-6	Threshold versus phase duration (pulse width) from Figure 2 of Pfingst et al. (1991), and deterministic and stochastic model predictions of these data for monopolar and bipolar stimulation. . . . .	109
6-7	Psychophysical data: threshold, uncomfortable loudness and the corresponding dynamic range versus phase duration. . . . .	110
6-8	Deterministic and stochastic model predictions of threshold and uncomfortable loudness and the corresponding dynamic range versus phase duration. . . . .	110
6-9	BP dynamic range divided by MP dynamic range—psychophysical data from White (Bruce, White, Irlicht, O’Leary and Clark, 1997) and model predictions, for 1, 2, 4 and 8 pulses. . . . .	112
6-10	BP dynamic range divided by MP dynamic range—psychophysical data from Battmer et al. (1993) and model predictions, for phase durations of 25, 50, 75, 100, 200 and 400 $\mu\text{s}$ /phase. . . . .	112
6-11	Effect of electrode configuration on psychophysical strength-duration curves. . . . .	114
6-12	Threshold versus number of pulses: psychophysical data and model predictions. . . . .	115

6-13	Intensity difference limen (Weber fractions) versus stimulus intensity (percentage of dynamic range): linear Weber function fits to psychophysical data from 8 subjects (dotted lines) and model predictions for $N_{\text{ucl}} = 100$ spikes (dot-dashed line), 500 spikes (dashed line) and 1 000 spikes (solid line). . . . .	115
6-14	Number of neural discharges versus stimulus intensity (percentage of dynamic range, with $N_{\text{ucl}} = 500$ spikes) predicted by the deterministic model (dashed line) and by the stochastic model (solid line) for a pulse rate of 125 pps. Error bars indicate $\pm 1$ std. . . . .	117
6-15	Deterministic and stochastic model predictions of dynamic range (left panes) and Weber fraction at 75% of dynamic range (right panes) versus threshold, for bipolar (top panes) and monopolar (bottom panes) stimulation, from models of 1 000 to 30 000 fibres, in response to a 125 pps pulse train. The number of fibres used in the model to generate each point are as labelled. Curves are plotted for two different values of $N_{\text{ucl}}$ , as indicated in the figure legend. . . . .	119
6-16	Linear fit to psychophysical data from Fig. 7 of Nelson et al. (1996): corresponding dynamic range (left pane) and Weber fraction at 75% of dynamic range (right pane) functions for a 125 pps pulse train. . . . .	119
7-1	Stochastic model of pulse-train response. . . . .	127
7-2	Schematic of pulse and bin numbering for calculation of the renewal time $r$ , where a discharge has occurred in bin $j$ of pulse 0 and the next subsequent discharge occurs in bin $i$ of pulse $n$ . A discretisation level of 6 bins per pulse is shown. . . . .	130
7-3	10 ms pulse train: Comparison of discharge rate mean and variance predicted by 500 iterations of Monte Carlo simulations for pulse rates of 200 pps ( $\Delta$ ) and 600 pps ( $\square$ ) and by the analytical approximation (solid lines). . . . .	135
7-4	20 ms pulse train: Comparison of discharge rate mean and variance predicted by 500 iterations of Monte Carlo simulations for pulse rates of 200 pps ( $\Delta$ ) and 600 pps ( $\square$ ) and by the analytical approximation (solid lines). . . . .	136
7-5	100 ms pulse train: Comparison of discharge rate mean and variance predicted by 500 iterations of Monte Carlo simulations for pulse rates of 200 pps ( $\Delta$ ) and 600 pps ( $\square$ ) and by the analytical approximation (solid lines). . . . .	137
7-6	Relative increase in threshold due to refractory effects: physiological data (o) and exponential least-squares fit (solid line). . . . .	139
7-7	Mean discharge rates for Neuron 3-21 of Javel et al. (1987) at pulse rates from 100 to 800 pps with a pulse width of 100 $\mu\text{s}$ /phase. Dashed lines indicate assumed extrapolations of the data. . . . .	140

7-8	Mean discharge rates for Neuron 3-21 of Javel et al. (1987) at pulse rates from 100 to 800 pps with a pulse width of 100 $\mu$ s/phase. Dotted lines are the first-pulse data of Neuron 3-21 scaled to indicate what the pulse-train response would be if inter-pulse interactions were not present. . . . .	141
7-9	Mean discharge rates for the deterministic version of the model at pulse rates of 100, 200, 300, 400, 600 and 800 pps with a pulse width of 100 $\mu$ s/phase. Dotted lines indicate mean discharge rates without refractory effects. . . . .	142
7-10	Mean discharge rates for the standard version of the stochastic model (i.e., <i>Case 1</i> ) at pulse rates of 100, 200, 300, 400, 600 and 800 pps with a pulse width of 100 $\mu$ s/phase. Dotted lines indicate mean discharge rates without refractory effects. . . . .	143
7-11	Mean discharge rates for the stochastic model with a modified refractory function (absolute refractory period = 1 ms; time-constant = 2 ms) at pulse rates of 100, 200, 300, 400, 600 and 800 pps with a pulse width of 100 $\mu$ s/phase. Dashed lines indicate mean discharge rates for the standard refractory function (from Figure 7-10).	144
7-12	Mean discharge rates for the <i>Case 2</i> version of the stochastic model at pulse rates of 100, 200, 300, 400, 600 and 800 pps with a pulse width of 100 $\mu$ s/phase. Dashed lines indicate mean discharge rates for the <i>Case 1</i> version (from Figure 7-10). . . .	145
7-13	Mean discharge rate data and model predictions for Neuron 3-10 of Javel et al. (1987) at pulse rates from 100 to 400 pps with a pulse width of 200 $\mu$ s/phase. . . .	146
7-14	Variance per pulse versus probability of discharge per pulse (o) for Neuron 3-21 of Javel et al. (1987) with a pulse width of 100 $\mu$ s/phase. Also plotted are the single-pulse stochastic model predictions (dotted lines) of these data, the stochastic model ( <i>Case 1</i> version) predictions (solid lines) and the deterministic model predictions (dashed lines). . . . .	148
7-15	Variance data (o) from Figure 7-14 are replotted, as are the stochastic model predictions of these data with the standard refractory function (dotted lines). Also plotted are the <i>Case 2</i> model predictions (dashed lines) and the stochastic model predictions with a modified refractory function (solid lines). . . . .	149
7-16	Variance per pulse versus probability of discharge per pulse (o) for Neuron 3-10 of Javel et al. (1987) with a pulse width of 100 $\mu$ s/phase. Also plotted are the single-pulse stochastic model predictions (dotted lines) of these data, the stochastic model ( <i>Case 1</i> version) predictions (solid lines) and the deterministic model predictions (dashed lines). . . . .	150

8-1 Discharge probability histograms for a model of 10 000 fibres in response to a 100 pps pulse-train with a bipolar electrode configuration, at intensities corresponding to  $N_{\text{uc1}} = 100$  and 500 spikes. Insets zoom in on the region of the histogram for discharge probabilities greater than zero. . . . . 161

# Preface

This thesis is the result of original research conducted by the author, except for contributions made by the thesis supervisors, Prof. G. M. Clark and Dr. L. S. Irlicht, and by co-authors of journal and conference papers arising from the research presented in this thesis. The publications resulting from each chapter and the contributions made by co-authors other than the thesis supervisors are as follows:

- Chapter 2: Some of the results of this chapter were published in a refereed paper presented at an international conference: Au, D., Bruce, I., Irlicht, L. and Clark, G. M. (1995). Cross-fiber interspike interval probability distribution in acoustic stimulation: A computer modelling study, *Ann. Otol. Rhinol. Laryngol.* **104**(Suppl. 166): 346–349. David Au and Laurence Irlicht developed the basilar membrane section of the computer model used in this chapter.
- Chapter 3: Some of the results of this chapter were published in a paper presented at an international conference: Bruce, I., Irlicht, L. and Clark, G. (1995). A mathematical analysis of sound coding in the auditory nerve: Spatiotemporal summing mechanisms, in P. P. Wang (ed.), *Proceedings of the Second Annual Joint Conference on Information Sciences*, Wrightsville Beach, NC, USA, pp. 296–299. A more complete description of the model derivation and results has been accepted for publication in an international peer-reviewed journal: Bruce, I. C., Irlicht, L. S. and Clark, G. M. (In Press). A mathematical analysis of spatiotemporal summation of auditory nerve firings, *Information Sciences: Applications*.
- Chapter 4: The results of this chapter were published in a paper presented at a national conference: Irlicht, L. S., Bruce, I. C. and Clark, G. M. (1996). A multi-threshold neural network for frequency estimation, *Proceedings of the Seventh Australian Conference on Neural Networks (ACNN'96)*, Australian National University, Canberra, Australia.
- Chapter 5: Some of the results of this chapter were presented at an international conference: Bruce, I. C., Irlicht, L. S., White, M. W., O'Leary, S. J., Dynes, S., Javel, E. and Clark, G. M.

(1997). A stochastic model of the electrically stimulated auditory nerve designed for the analysis of large-scale population response, *ARO Midwinter Meeting Abstracts*. A fuller description of the model derivation and results of this chapter have been submitted to an international peer-reviewed journal: Bruce, I. C., Irlicht, L. S., White, M. W., O’Leary, S. J., Dynes, S., Javel, E. and Clark, G. M. A stochastic model of the electrically stimulated auditory nerve: Single-pulse response, *Submitted to IEEE Trans. Biomed. Eng.* Mark White assisted in the development of the modelling method, and Stephen O’Leary performed some of the analysis of the physiological data which was contributed by Scott Dynes and Eric Javel, as noted in the text.

Chapter 6: Some of the results of this chapter were presented as a poster at an international conference: Bruce, I. C., Irlicht, L. S., White, M. W., O’Leary, S. J. and Clark, G. M. (1997) Electrical stimulation of the auditory nerve: Prediction of psychophysical performance by a model including stochastic aspects of neural response, *ARO Midwinter Meeting Abstracts*; others were published in a paper presented at another international conference: Bruce, I. C., Irlicht, L. S., White, M. W., O’Leary, S. J., Dynes, S., Javel, E. and Clark, G. M. (1997). An improved model of electrical stimulation of the auditory nerve, in G. M. Clark (ed.), *Cochlear Implants. XVI World Congress of Otorhinolaryngology Head and Neck Surgery*, Monduzzi Editore, Bologna, pp. 125–130. A more complete description of the modelling methods and results has been submitted to an international peer-reviewed journal: Bruce, I. C., White, M. W., Irlicht, L. S., O’Leary, S. J. and Clark, G. M. The effects of stochastic neural activity in a model predicting loudness perception with cochlear implants: Low-rate stimulation, *Submitted to IEEE Trans. Biomed. Eng.* Mark White and Stephen O’Leary assisted in the development of the modelling method. Mark White also contributed some psychophysical data, as noted in the text.

Chapter 7: Some of the results of this chapter were presented at an international conference: Bruce, I. C., Irlicht, L. S., White, M. W., O’Leary, S. J., Dynes, S., Javel, E. and Clark, G. M. (1997). A stochastic model of the electrically stimulated auditory nerve designed for the analysis of large-scale population response, *ARO Midwinter Meeting Abstracts*. A fuller description of the model derivation and results of this chapter have been submitted to an international peer-reviewed journal: Bruce, I. C., Irlicht, L. S., White, M. W., O’Leary, S. J., Dynes, S., Javel, E. and Clark, G. M. A stochastic model of the electrically stimulated auditory nerve: Pulse-train response, *Submitted to IEEE Trans. Biomed. Eng.* Mark White assisted in the development of the modelling method, and Stephen O’Leary performed some of the analysis of the physiological data which was contributed by Scott Dynes and Eric Javel, as noted in the text.

# Chapter 1

## Introduction

In the human auditory system, the function of the ear is to transduce sound waves into spatiotemporal patterns of activity in the auditory nerve, which are processed by the brain to produce the perception of sound. *Spatial* refers to the fact that the auditory nerve consists of some 30 000 nerve fibres, each driven by vibrations at different sites in the cochlea such that the spread of neural excitation is a function of the acoustical stimulus. *Temporal* refers to the fact that the timing of discharges in each fibre is also a function of the stimulus. This transduction process can be described as the *coding* of sound features in the patterns of activity in the auditory nerve. This neural code is propagated to the brainstem by the auditory nerve and processed (decoded) by the auditory pathways in the brain. A number of different forms of neural processing or decoding may take place to form the perception of sound, or hearing.

Sensorineural hearing loss refers to a disruption in the transduction of sound waves into neural activity in the auditory nerve as a result of malformation or degeneration of the organ of Corti within the inner ear, or *cochlea*. In the case of a severe hearing loss, the disruption is of such a degree that little perception of sound can be achieved, even at high sound levels. With a profound hearing loss this disruption is practically complete. In both cases, amplification of sound by a hearing aid will provide very little assistance. However, because of the electrochemical nature of auditory nerve activity, an alternative exists—electrical stimulation of the nerve by a device known as a *cochlear implant* or *hearing prosthesis* can be used to produce neural discharges of the auditory nerve. Consequently, cochlear implants provide a means of circumventing the disruption to auditory processing resulting from a sensorineural hearing loss, by electrically processing sound waves and transducing them via electrical stimulation into activity in the auditory nerve.

If it were possible to design a cochlear implant which perfectly replicated the function of the human ear, then it would not be necessary to understand how the brain interprets patterns of activity in the auditory nerve produced by the electrical stimulation in order to develop that implant—the implant could simply mimic the action of a healthy ear. However, present cochlear

implant technology is unable to perform such advanced processing of sound as the human ear and is limited in what patterns of neural activity it can elicit. With such sub-optimal coding of sound it is desirable to produce neural responses which maximally convey the most *perceptually-important* information. Thus, in developing cochlear implants, it is necessary to understand both (i) how the brain interprets patterns of auditory nerve activity and (ii) what patterns of response are produced by electrical stimulation. The general aim of this thesis is to develop mathematical and computational models to:

1. determine which components of auditory nerve activity may be significant for aspects of perception such as threshold, loudness and pitch, and
2. explain both the electrically elicited auditory nerve activity (physiological responses) and the resulting perception of the stimulus (psychophysical performance) in cochlear implant users.

The methodology used is first to derive mathematical descriptions of auditory nerve responses to acoustical and electrical stimulation from physiological data and theory, and then to apply statistical measures of how stimulus features are represented in different schemes of spatiotemporal combination and decoding of patterns of activity. Applications of these schemes to perceptual phenomena are considered theoretically, and some comparisons of predicted responses are made to those obtained from psychophysical experiments.

Before moving on to an overview of the specific issues addressed in this thesis, a brief summary is provided of the anatomy and physiology of the ear, how the normal ear may be coding sound features, and how cochlear implants function.

## 1.1 Anatomy and physiology of the ear

A detailed description of the anatomy and physiology of the ear can be obtained from standard texts such as Pickles (1982). A more recent review of the cochlea is given in Dallos et al. (1996). This section constitutes a summary of the anatomical and physiological features which are pertinent to this thesis.

The anatomy of the peripheral auditory system, as shown in Figure 1-1, is generally divided into three segments: the outer, middle, and inner ears. The outer ear consists of the pinna (**A**) and the external auditory meatus (**B**), which function to capture incident sound waves and provide initial filtering of the signal, increasing the sound pressure in the region of 2 to 7 kHz and assisting in perceptual tasks such as sound localisation.

The middle ear comprises the tympanic membrane (**C**) and three ossicles—the malleus (**D**), the incus (**E**) and the stapes (**F**)—which function to couple the external auditory meatus to the cochlea (**H**). Sound pressure waves in the auditory meatus produce vibrations of the tympanic

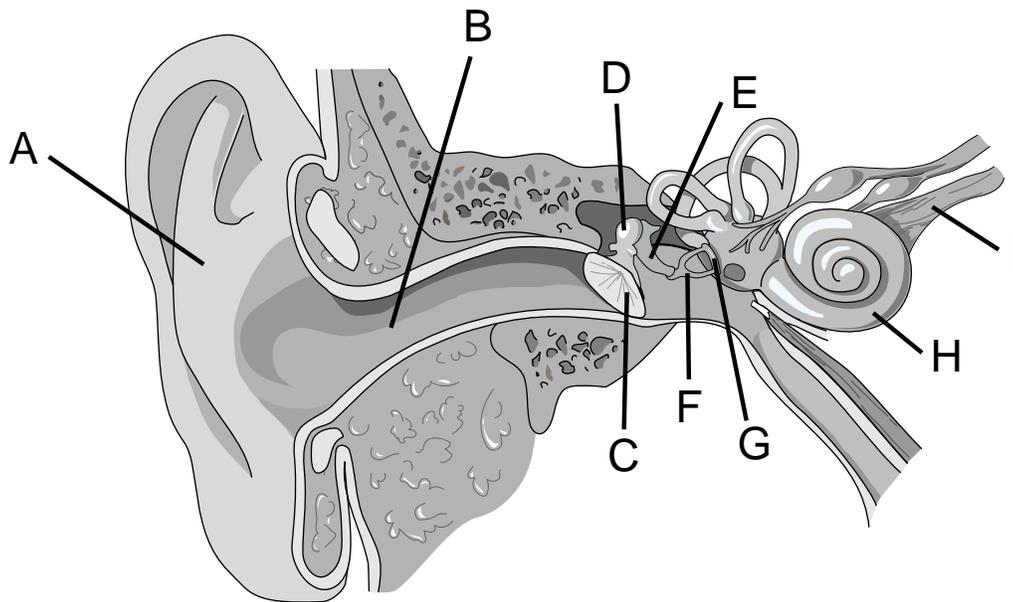


Figure 1-1: Anatomy of the human ear: (A) pinna; (B) external auditory meatus; (C) tympanic membrane; (D) malleus; (E) incus; (F) stapes; (G) oval window; (H) cochlea; (I) auditory nerve.

membrane which are transferred by the ossicles to the cochlea. The middle ear's mechanics are such that they match the impedance of the auditory meatus to the much higher impedance of the cochlear fluids, increasing transfer of sound energy from the outer to the inner ear and reducing reflection of sound waves at the tympanic membrane. The first two ossicles are joined relatively rigidly so that when the tip of the malleus is pushed by the tympanic membrane, the ossicles rotate together and transfer the force to the stapes. The footplate of the stapes is attached to the oval window (G), a flexible window in the wall of the cochlea.

The cochlea is longitudinally divided into three fluid-filled scalae (the scala vestibuli, the scala tympani, and the scala media) as shown schematically in Figure 1-2. The scalae spiral together along the length of the cochlea, with the outer scalae (the scala vestibuli and the scala tympani) joined by an opening at the apex of the cochlea, the helicotrema. The scala media forms a closed inner compartment, which is separated from the scala vestibuli by Reissner's membrane on one side and from the scala tympani by the basilar membrane on the other side. Movement of the oval window, situated at the base of the scala vestibuli, produces a displacement of the cochlear fluids to a second window in the cochlea, the round window, which opens into the base of the scala tympani. The flow of the cochlea fluids results in a wavelike displacement of the basilar membrane, referred to as the *travelling wave*. This vibration of the cochlear membranes is indicated by the dashed lines in Figure 1-2.

The travelling wave propagates from the base of the basilar membrane to its apex. The rigidity

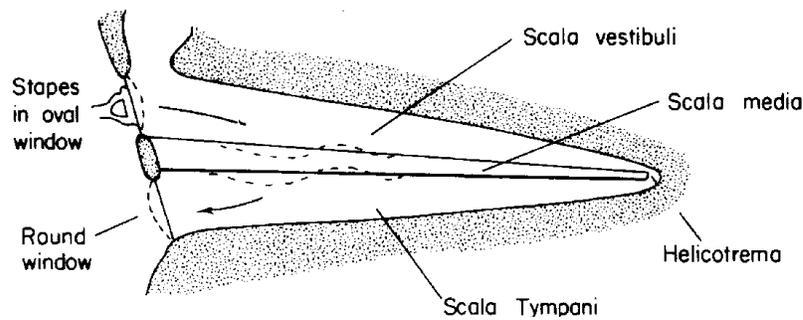


Figure 1-2: Schematic of the vibrations in the cochlear duct, which is depicted as if it had been unrolled. From Pickles (1982).

of the basilar membrane decreases with distance from the base, and consequently the wave amplitude grows and the propagation velocity drops as it moves apically until a region is reached where the cochlear partition approaches resonance and the wave starts to stall. Following this region of resonance the vibration amplitude falls dramatically and the travelling wave propagates no further. The site of the resonance is dependent on the frequency of stimulus, such that different frequencies produce different areas of maximum vibration—this is referred to as a *tonotopic* organisation. Consequently, the basilar membrane can roughly be described as a continuous series of band-pass filters, with the centre frequency of each filter decreasing monotonically with distance from the base. The reduction in propagation velocity in the region of resonance also produces a rapid shift in the phase of the vibration, leading to a phase lag which accumulates as the wave increases in amplitude to its maximum and then dies off.

A cross-section of the cochlea is shown in Figure 1-3. The organ of Corti is a structure which sits on the basilar membrane and performs (i) active filtering of the basilar membrane vibration and (ii) the transduction of the sound energy conveyed by the travelling wave into neural activity within the auditory nerve. The former is achieved by the *outer* hair cells contained within the organ of Corti—details of the organ of Corti are shown in Figure 1-4. Hairs, or *stereocilia*, extend from the outer hair cells on the side of the organ of Corti opposite the basilar membrane and are embedded in the tectorial membrane, which rests on that side of the organ of Corti. The tectorial membrane is only attached on one side, whereas the basilar membrane is suspended between the spiral lamina and the spiral ligament. Consequently, displacement of the basilar membrane in the form of the travelling wave causes a shear movement between the organ of Corti and the tectorial membrane, forcing the stereocilia to bend. Slender filaments form links between the tops of short stereocilia and the sides of adjacent tall stereocilia. It is hypothesised that changes in the tension of these tip links, caused by bending of the stereocilia, produce changes in the motility of the outer hair cells, actively modifying the basilar membrane vibration and consequently increasing its sensitivity and frequency selectivity. This active response may also be modulated by an efferent

nerve supply which the outer hair cells receive from the cochlear nucleus.

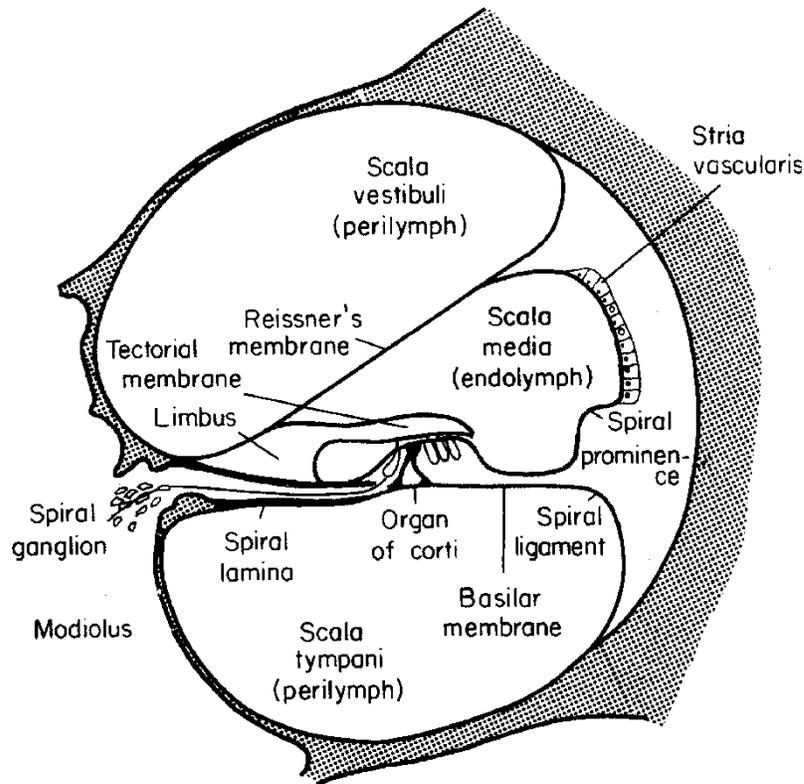


Figure 1-3: A cross-section of the cochlea. From Pickles (1982).

The transduction of the sound energy conveyed by the travelling wave into neural activity within the auditory nerve is achieved by the *inner* hair cells and auditory nerve synapses which are also contained within the organ of Corti. The stereocilia of the inner hair cells extend towards the tectorial membrane, like those of the outer hair cells, but they do not appear to be embedded in it. However, the shear movement of the tectorial membrane still forces the stereocilia of the inner hair cells to bend, either because they are loosely attached to the tectorial membrane, or because the cochlear fluid between the organ of Corti and the tectorial membrane is made to flow around the stereocilia by the shear movement, exerting a lateral force on them. Bending of the stereocilia produced by the basilar membrane vibration can lead to current flow within the hair cells, which causes the release of neurotransmitters across the hair-cell/auditory-nerve synapses.

Auditory nerve fibres provide a direct synaptic link between the inner hair cells and neurons in the auditory brainstem, or *cochlear nucleus*. Around 95% of all afferent auditory nerve fibres synapse onto inner hair cells, with 10 to 30 fibres each forming a single synaptic connection to an inner hair cell. A fibre's axons and cell body are effectively a continuous tubular myelinated membrane, with regular breaks in the myelin referred to as the nodes of Ranvier. Ionic fluids fill and surround the nerve membrane, with differing concentrations of ions in the internal and external

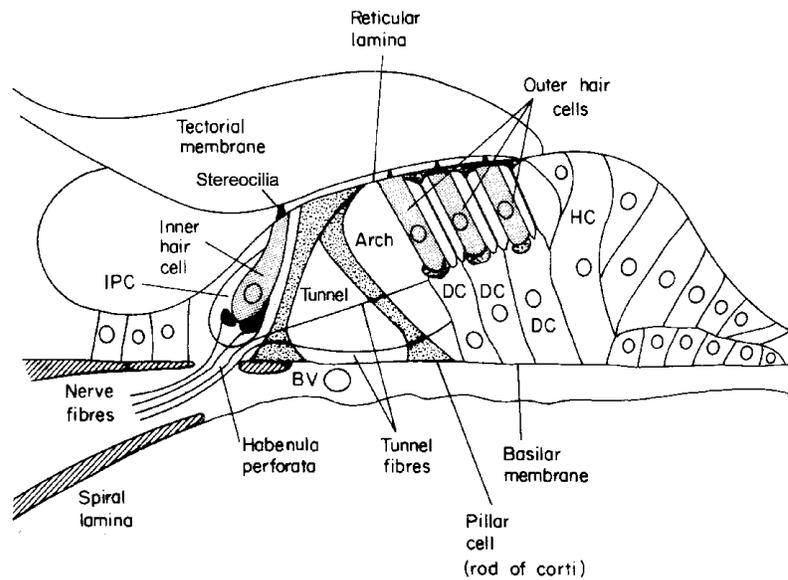


Figure 1-4: A cross-section of the organ of Corti. From Pickles (1982).

fluid. The electrical charges of the ions are such that there exists a difference in potential between the internal and the external fluids. The reception of the neurotransmitters by an auditory nerve synapse causes ions to flow through the membrane from the interior to the exterior of the fibre and vice versa. This results in a change in the potential of a nerve membrane, known as an excitatory postsynaptic potential. If the postsynaptic potential is of sufficient magnitude, it is propagated along the auditory nerve fibre to the brain. Such a response is called an *action potential*, *discharge*, *firing* or *spike*. This neural activity is then processed by the brain to give the perception of sound.

Inner hair cells appear to release small amounts of neurotransmitters even when they are in their resting state, i.e., in the absence of basilar membrane vibration. This leads to auditory nerve discharges occurring in the absence of an acoustical stimulus, or *spontaneous activity*, with the rate of spontaneous discharges in single fibres ranging from close to zero up to 100 spikes per second. In the presence of a weak acoustical stimulus, the basilar membrane vibrates and the stereocilia bend back and forth, as described above. This produces an oscillation in the amounts of neurotransmitters released, such that action potentials are more likely to occur at one phase of the stimulus than at others. This is known as *phase-locking* or *synchronisation*, and can occur even though there may be no net increase in the average rate of firing. As the magnitude of the basilar membrane vibration increases, the amounts of neurotransmitters released also increase, such that both the average discharge rate and the synchronised component of the discharge rate are roughly monotonic functions of the vibration amplitude. The discharge rate does not increase linearly however, because of (i) nonlinearities in the transduction process and (ii) refractory effects which limit the rate of recovery of a fibre after it has discharged to a state where it can discharge again.

This limits the maximum average discharge rate to around 1 000 spikes per second. However, this is only achieved for very brief stimuli or at the onset of longer continuous stimuli—adaptation in the transduction process causes an exponential decrease in the discharge rate over the duration of a stimulus, asymptoting to a maximum discharge rate of around 200–300 spikes per second. Auditory nerve fibres exhibit a range of thresholds (typically defined as the intensity at which the average discharge rate rises above its spontaneous rate) on the order of 60–80 dB<sup>1</sup>, although perhaps 70% of fibres fall within the bottom 10 dB. Dynamic ranges (the increase in stimulus magnitude required to increase the average discharge rate from its threshold to saturation) are on the order of 20–50 dB. The above description of auditory nerve response as a function of the basilar membrane vibration provides an overview of the *temporal* response characteristics of auditory nerve fibres. However, as discussed previously, the behaviour of these vibrations is different at each point along the basilar membrane, giving rise to a *spatial* pattern of responses also. The tonotopic arrangement of vibration patterns and the delays in the vibration phase are reflected in the spread of excitation in the auditory nerve and in the phase of the synchronised components of response in different fibres respectively. Consequently, patterns of auditory nerve activity are described in terms of their spatiotemporal characteristics.

Such a spatiotemporal pattern of activity is illustrated in Figure 1-5, which shows the output of a computer simulation<sup>2</sup> in response to a complex acoustical stimulus. The instantaneous probability of firing for 19 auditory nerve fibres is plotted as a function of their origin on the basilar membrane (distance from the stapes) and of time. In any one fibre the response is seen to change over time, and at any one time the response of each fibre is seen to be different. The basilar membrane displacement at two different times (dashed and solid lines) is plotted on the left, indicating the movement of the travelling wave which produces this neural response.

## 1.2 Coding of sound in the normal ear

Although physiological, psychophysical and modelling studies have revealed much about which aspects of auditory nerve activity are likely to be coding different sound features, many issues remain unresolved. Therefore, reviewing all issues relating to sound coding is beyond the scope of this chapter—a recent review of physiological models for basic auditory percepts can be found in Delgutte (1996). Instead, only a summary of the general principles of sound coding theories is provided here, and the specific issues relating to coding of sound by cochlear implants will be addressed later in the introduction to cochlear implants (Section 1.3) and the thesis overview (Section 1.4).

The spatiotemporal nature of the auditory nerve activity means that information may be coded

---

<sup>1</sup>All values given in units of dB are calculated by  $20 \log_{10}(\text{ratio})$ , unless otherwise stated.

<sup>2</sup>See Chapter 2 for details of the computer model.

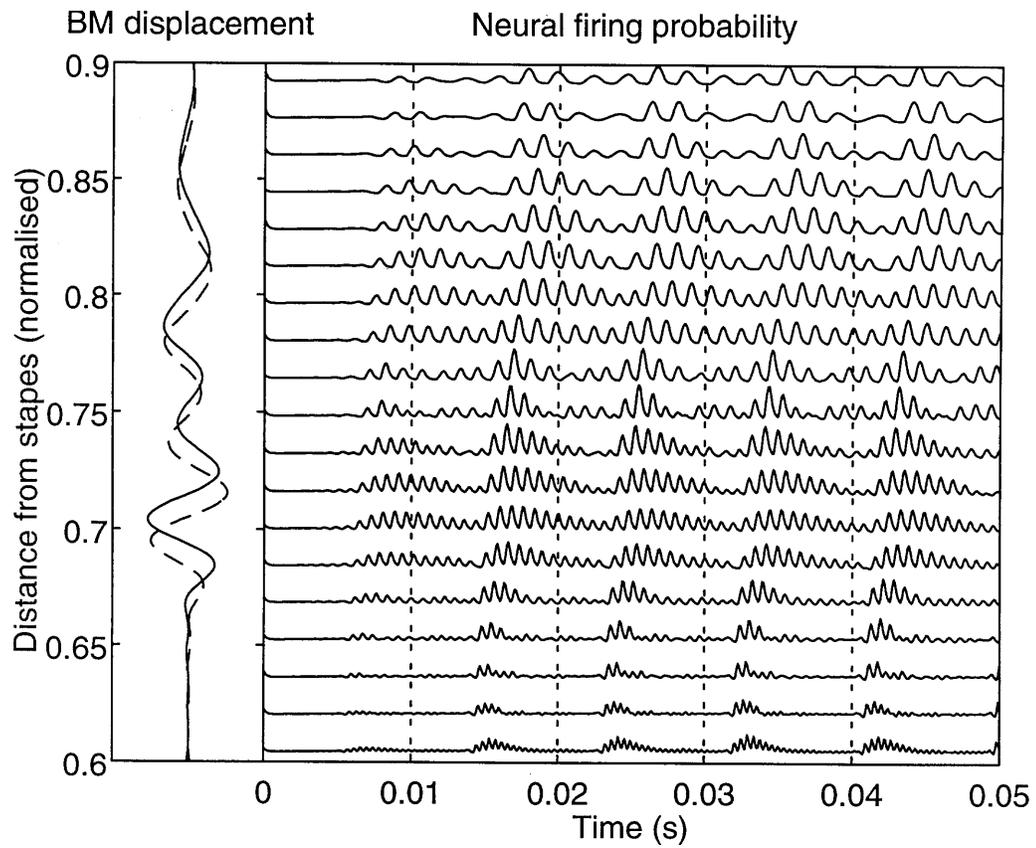


Figure 1-5: Instantaneous probability of firing (indicated by the height of the trace) for 19 auditory nerve fibres versus origin on the basilar membrane (distance from the stapes) and time. The basilar membrane displacement at two different times (dashed and solid lines) is shown on the left, indicating the movement of the travelling wave which produces this neural response.

in the site of the activity (spatial coding), in the timing of the activity (temporal coding), or in some combination of the two (spatiotemporal coding).

### 1.2.1 Spatial coding

Theories of spatial coding propose that acoustical stimulus features are coded by *which* auditory nerve fibres respond. The fibres which are deemed to be ‘responding’ may be distinguished from ‘non-responding’ fibres by their response properties.

For example, the frequency of an acoustical stimulus may be coded by which fibres exhibit the maximum mean discharge rate (for a review, see Evans, 1981). However, such coding may be limited by restricted dynamic ranges in single fibres and overlap of the neural populations excited by different frequency components of complex stimuli at high intensities (e.g., Evans, 1981). Variations to this coding scheme which may overcome such restrictions include: frequency is coded by the fibres exhibiting the greatest change in mean discharge rate *with respect to their spontaneous rate* (e.g., Kim et al., 1990), and frequency is coded by the fibres exhibiting the *greatest difference in mean discharge rate across a local population of fibres* (Kim and Parham, 1991).

Alternatives have also been proposed as to what response property is used to distinguish fibres for spatial frequency coding. It has been found that the *degree of synchronisation in a local population of fibres* (Young and Sachs, 1979) well represents the frequency components of complex stimuli even at high intensities (e.g., Jenison et al., 1991; Young and Sachs, 1979). A neural mechanism which may utilise such coding is investigated in Chapter 3. Another possibility is that the fibres exhibiting the *greatest change in phase across a local population of fibres* code the stimulus frequency (Shamma, 1985).

Not only the stimulus frequency may be coded by the spatial patterns of response. One long-standing theory of intensity coding is that the stimulus intensity is coded by the spread of neural excitation elicited by an acoustical stimulus (e.g., Moore and Glasberg, 1996; Zwicker and Fastl, 1990; Zwicker and Scharf, 1965; Zwicker, 1958).

### 1.2.2 Temporal coding

In theories of temporal coding, it is hypothesised that information about an acoustical stimulus is coded by the *timing* of auditory nerve fibre responses.

For instance, one theory of *temporal* frequency coding which competes with theories of *spatial* frequency coding is that frequency is coded by interspike intervals in single fibres or in populations of fibres (e.g., Clark, Carter, Maffi and Shepherd, 1995; Rose et al., 1967; Wever, 1949). However, it is generally recognised that such coding is probably limited to stimulus frequencies lower than a few thousand Hertz and that spatial frequency coding is required for higher stimulus frequencies. Neural mechanisms which may utilise temporal frequency coding are investigated in Chapters 2

to 4.

Additionally, the stimulus intensity may be coded not just by the spread of excitation, but rather by the average discharge rates in single fibres or in populations of fibres (e.g., Winter and Palmer, 1991; Hellman and Hellman, 1990; Viemeister, 1988; Viemeister, 1983; Evans, 1981). A neural mechanism which may utilise such coding is investigated in Chapter 3, and a psychophysical model which utilises such coding is explored in Chapter 6. One hypothesised temporal coding scheme which overcomes limits in average-discharge-rate intensity coding is that the *degree of coincidence of discharges in a population of fibres* is used to code intensity (Carney, 1994).

### 1.2.3 Spatiotemporal coding

Although theories of coding are typically classified under the headings of spatial coding or temporal coding, almost all of the coding schemes listed above actually involve coding of stimulus features in the temporal features of a population of fibres, i.e., *spatiotemporal* coding. In theories of ‘spatial’ coding, temporal features of activity in populations of fibres are used to indicate which fibres are responding. In ‘temporal’ coding schemes, multiple fibres are typically required, either because physiological data have suggested that certain aspects of neural behaviour, such as refractory effects, limit what information can be coded in single auditory nerve fibres (e.g., Clark, Carter, Maffi and Shepherd, 1995; Rose et al., 1967; Wever, 1949), or because the theoretical performance of a detection, discrimination or estimation system based on single-fibre responses can not be reconciled easily with psychophysical data (e.g., Winter and Palmer, 1991; Hellman and Hellman, 1990; Viemeister, 1988; Viemeister, 1983; Evans, 1981).

## 1.3 Cochlear implants

A sensorineural hearing loss leads to significant, or in some cases complete, reduction in the number and integrity of inner hair cells (Hinojosa and Marion, 1983). Consequently, very little or no spontaneous auditory nerve activity is present in a severely to profoundly deafened ear (Kiang et al., 1970), and little or no neural activity is produced by an acoustical stimulus, which results in a hearing impairment. Retrograde degeneration of the auditory nerve also occurs, reducing the number of surviving auditory nerve fibres (Hinojosa and Marion, 1983). However, in most individuals, enough fibres remain to enable electrical stimulation via a cochlear implant to elicit an audible degree of neural activity. An electronic processor can be used to extract features of acoustical signals and to transfer them into spatiotemporal patterns of electrical current which will produce appropriate neural responses. Such processors are typically optimised for speech signals and are therefore often referred to as *speech processors*.

As shown in Figure 1-6, a typical cochlear implant consists of a microphone (**a**), a speech

processor—worn either behind the ear (**b**) or on the body (**c**)—and an electrode array (**f**). Either a transcutaneous or a percutaneous link is used to connect the electrode array to the speech processor. The implant system shown here uses a transcutaneous RF transmitter coil (**d**) and receiver/stimulator (**e**) to communicate between the processor and the stimulator and to provide the stimulator's operating electrical power. The present generation of cochlear implants utilise an array of up to 24 electrodes which is inserted into the basal turns of the scala tympani of the cochlea (**g**) through the round window or via a cochleostomy. Some implants have additional electrodes external to the cochlea. The creation of a potential difference between two or more electrodes on the electrode array (**f**) by the stimulator (**e**) causes electrical currents to flow in the cochlea (**g**) and auditory nerve (**h**). The two electrode configurations investigated in this thesis are commonly known as *monopolar* and *bipolar* stimulation. In the case of monopolar stimulation, the active electrode is one of the electrodes on the array within the cochlea and the return electrode is an electrode external to the cochlea. In the case of bipolar stimulation, both the active electrode and the return electrode are on the electrode array within the cochlea. The electrode configuration will be referred to as being *narrow bipolar* if the active and return electrodes are adjacent to each other on the array, and *wide bipolar* if there are nonactive electrodes separating the two.

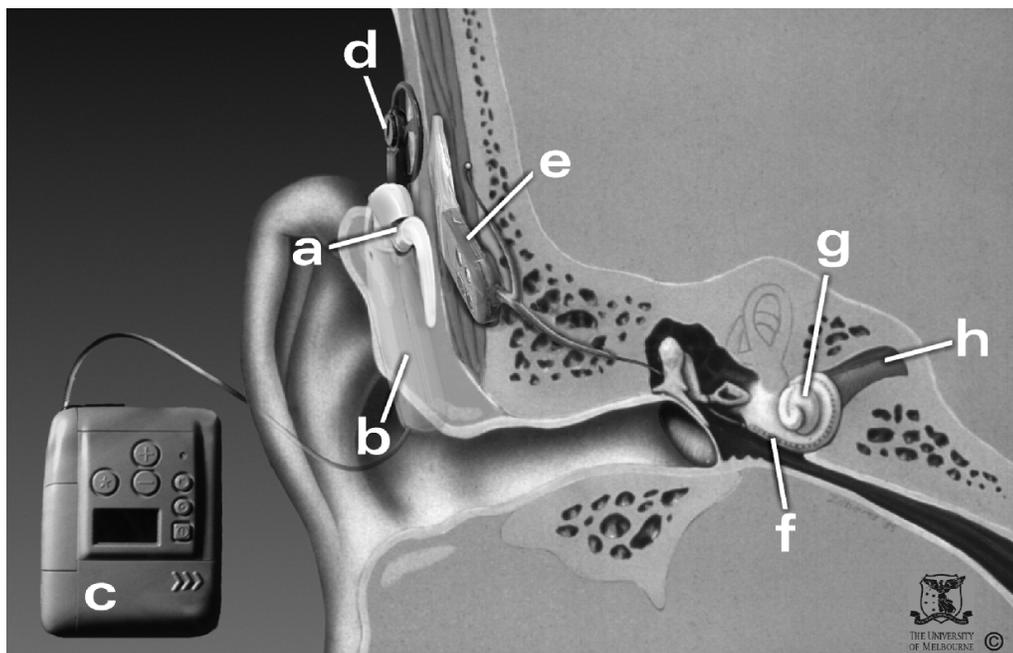


Figure 1-6: Components of a typical cochlear implant: (a) microphone; (b) behind-the-ear speech processor; (c) body-worn speech processor; (d) transmitter coil; (e) receiver/stimulator; (f) electrode array; (g) cochlea; (h) auditory nerve.

Injection of current by means of a cochlear implant can cause a sudden depolarisation in the nerve membrane, similar to an excitatory postsynaptic potential produced by a healthy hair-

cell/auditory-nerve synapse (Frijns, 1995; Colombo and Parkins, 1987). The magnitude of this depolarisation is dependent on the geometry of the current flow relative to the nerve membrane and on the state of the membrane (Frijns, 1995; Ranck, 1975). If the depolarisation is of sufficient magnitude, it will form a self-sustaining depolarisation which propagates the length of the fibre to the cochlear nucleus (Frijns, 1995; Colombo and Parkins, 1987). Such discharges are practically the same as action potentials generated by a healthy ear and will produce sensations perceived as sounds by cochlear implant users. Which fibres respond will depend on the stimulus current produced at the active electrode, how the current flows to surviving auditory nerve fibres, and how effectively the membrane is depolarised by this current. In general, fibres closer to the active electrode are more easily excited (van den Honert and Stypulkowski, 1987a; Ranck, 1975), and the range of thresholds in a local population of fibres is on the order of 10–14 dB (van den Honert and Stypulkowski, 1987a). Multiphasic (particularly biphasic), charge-balanced current pulses are typically used to avoid damaging charge build-up in the neural tissue (for a recent review, see Tykocinski et al., 1995b). Stimulation on multiple electrodes is normally required to be non-simultaneous, because field interactions created by simultaneous stimulation can cause unpredictable and undesirable neural responses (Clark, 1987; White et al., 1984; Eddington et al., 1978). Thresholds for the cathodic phase are generally lower than for the anodic phase (Ranck, 1975; Hill, 1936). In a single fibre, a 1–6 dB increase in the stimulus intensity is typically required to increase the discharge probability in response to a single current pulse from close to zero up to nearly one (see Chapter 5). Discharges tend to be very well phase-locked to the leading edge of a current pulse, with only a small amount of jitter occurring in the timing of the discharge (Javel et al., 1987; van den Honert and Stypulkowski, 1987b; van den Honert and Stypulkowski, 1984). Responses to trains of current pulses exhibit inter-pulse interactions such as refractory effects, which alter the temporal pattern of activity (see Chapter 7).

A number of dissimilarities exist between neural response patterns elicited by a healthy ear and those produced by a cochlear implant. These contribute to deficiencies in the speech perception of implant users, which is still below that of individuals with normal hearing (e.g., Clark, 1997a). The ability of a particular implant and speech processing strategy to produce the desired neural response is restricted by a number of factors. Firstly, the software and hardware limitations of the speech processor, implant and electrode array restrict what electrical currents may be produced via the electrodes. Secondly, the electrical field produced at sites of action potential generation in the auditory nerve will be affected by the behaviour of those electrical fields within the structures of the electrode array, cochlea and auditory nerve (Frijns, 1995; Finley et al., 1990). Thirdly, action potentials in the normal cochlea are generated by the mechanical/electrical transduction action of hair cells, not by direct electrical fields such as those produced by cochlear implants. Even if the first two restrictions did not exist, and any desired pattern of stimulating current were

achievable, this difference in the way in which action potentials are generated could mean that not all patterns of neural response are achievable via electrical stimulation with present electrode designs. Consequently, auditory nerve firing patterns produced by present cochlear implants are very different to those of normal hearing. It is therefore necessary to optimise the coding of sound features which are the most important for perception.

The first successful cochlear implants consisted of a single stimulating electrode, limiting the coding of speech features to the temporal aspects of a single analogue or pulsatile current (House and Berliner, 1982; Hochmair-Desoyer et al., 1980). Processing strategies for single-electrode implants typically encoded the broadband envelope of an acoustical stimulus in the amplitude envelope of the analogue or pulsatile current (House and Berliner, 1982; Hochmair-Desoyer et al., 1980).

The development of multi-electrode implants made possible stimulation of different populations of auditory nerve fibres (e.g., Clark, 1987; Eddington, 1980), permitting the coding of more than just the stimulus envelope. The next most important speech features were deemed to be the formants and higher frequency components (Clark, 1987). As described previously, both the site of neural excitation and the periodicity of firing in single fibres change as a function of acoustical stimulus frequency. Consequently, most theories of frequency coding are based on either of these two features of auditory nerve activity (for a recent review, see Delgutte, 1996). Preliminary psychophysical investigations with a multi-electrode implant (Clark, 1987) indicated that either the site of stimulation or the rate of the stimulating pulse-train could be varied to produce a change in pitch, i.e., the psychological percept of frequency. Both coding schemes have been implemented in speech processors for multi-electrode cochlear implants. One approach, utilising temporal (firing periodicity) frequency coding, was to code speech features such as frequency of the voicing pitch and formants by the periodicity of pulses in multiple non-simultaneous, pulsatile currents (Clark, 1987). Another approach, utilising spatial (site of excitation) frequency coding, was to code spectral components of speech by continuous spectral analysis by a fixed number of filters and continuous analogue output from electrodes exciting different neural populations (Eddington, 1980). Most theories of intensity coding are based on degree of neural excitation (Evans, 1981), and neural excitation is known to increase with the current delivered by an electrical stimulus. Further psychophysical investigations (Clark, 1987) indicated that varying the current per pulse (amplitude and/or pulse-duration) produces a change in loudness, i.e., the psychological percept of intensity. Consequently the intensity of speech features or spectral components was coded by the amount of current delivered by the corresponding electrode.

Speech perception with multi-electrode implants was better on average than with single electrode devices (Gantz et al., 1987). The multi-electrode ‘temporal coding’ processing strategy was later superseded by a spectral-analysis strategy—speech perception was found to improve on average by using a processor which chooses the position of the stimulating electrodes according to

the natural tonotopic arrangement of the auditory nerve to represent the spectral components of the speech signal with greatest amplitude (Clark, Dowell, Cowan, Pyman and Webb, 1995; McKay and McDermott, 1993; McKay et al., 1991). This was executed with non-simultaneous, pulsatile stimulation, but at a constant rate, i.e., with no coding via timing of pulses. The ‘spatial coding’ multi-electrode strategy with a fixed number of filters and electrodes has also been adapted to make use of non-simultaneous, pulsatile stimulation rather than continuous analogue stimulation (Wilson et al., 1993; Merzenich and White, 1977). With improvements in the software and hardware used in speech processors and implants, it has been possible to stimulate at higher rates than initially was possible. This can produce improvements in speech perception (e.g., Wilson et al., 1991), although increasing the pulse rate above 800 pulses per second with some strategies can cause performance to decline (Zilberman, 1997).

This history of cochlear implant development displays a move away from coding speech features via the temporal aspects (stimulation waveform or timing of pulses) of the electrical stimulus to coding spectral features via the spatial aspects (stimulating-electrode position). However, speech perception with a *fixed-rate* stimulation scheme may be limited, and more precise spatiotemporal patterns of response may be needed to further improve the performance of speech processing strategies. For instance, some aspects of pulse-timing utilised in earlier speech processing strategies to code speech features, or some new aspects such as mimicking the basilar membrane phase lag behaviour (Clark, 1997b; Clark, 1996; Clark, Carter, Maffi and Shepherd, 1995), if introduced into coding schemes which also provide a high level of spectral information, may further improve speech perception. However, despite the success of the initial ‘temporal coding’ scheme, it is still not fully understood how such temporal aspects of neural firing are coding stimulus features. It is therefore desirable to gain a better understanding of this as attempts are made to introduce aspects of temporal coding into speech processing strategies.

## 1.4 Thesis overview

Single fibres are limited in the information they can code, both in their long-term average response (e.g., Evans, 1981; Siebert, 1970) and in the periodicity of their activity (e.g., Clark, Carter, Maffi and Shepherd, 1995; Johnson, 1980; Siebert, 1970; Rose et al., 1967). Therefore, it has been proposed that activity in groups of auditory nerve fibres may be combined at different stages of the auditory pathways (Clark, 1996; Rothman and Young, 1996; Carney, 1994; Rothman et al., 1993; Carney, 1992; Winter and Palmer, 1991; Kim and Parham, 1991; Kim et al., 1990; Viemeister, 1988; Rhode and Smith, 1986; Teich and Lachs, 1979; Rose et al., 1967; Wever, 1949). Much of this combining may occur in the convergence of auditory nerve fibres onto neurons in the cochlear nucleus (Paolini et al., 1997; Joris, Carney, Smith and Yin, 1994; Joris, Smith and Yin, 1994; Young

et al., 1988; Smith and Rhode, 1987), i.e., direct combination of the activity in groups of auditory nerve fibres. In Chapter 2 a possible mechanism is investigated for temporal frequency coding via spatiotemporal combination of activity in pairs of auditory nerve fibres. In Chapter 3 this investigation is extended to look at how another neural mechanism, spatiotemporal summation, could be used in spatiotemporal combination of activity in more than two fibres. In these first two chapters a model of auditory nerve response to acoustical stimulation is used which has identical fibres, and only responses to pure-tone stimuli are investigated. However, the response characteristics of auditory nerve fibres do vary somewhat throughout the auditory nerve, particularly in how they respond to complex sounds. Therefore, in Chapter 4 a model is developed to examine how the combination of activity from fibres with different thresholds may assist in the decoding of different components of complex acoustical stimuli.

In the case of electrical stimulation, it has generally been assumed that the auditory nerve fibres have a deterministic (non-random) response—i.e., neural responses are identical for repeated identical stimuli—despite significant physiological evidence for stochastic (random) activity at some stimulus intensities. Consequently, it has been assumed that the temporal pattern of neural response closely matches the timing of the stimulating electrical pulses, except where refractory effects prevent a fibre from discharging. However, if cochlear implants are operating at stimulus intensities which produce stochastic activity in the auditory nerve response, this will significantly alter the spatiotemporal patterns of response to electrical stimulation. Furthermore, deterministic models (Pfungst, 1990; Shannon, 1989) and analyses of single-fibre physiological data which assume a deterministic response (Pfungst et al., 1991; Pfingst, 1990; Pfingst, 1988) have been unable to explain some important psychophysical phenomena observed in cochlear implant users. Therefore, in Chapter 5 the physiological evidence for stochastic activity in the auditory nerve response to electrical stimulation is reviewed and a model is developed which accurately describes the physiological data. In order to determine whether or not inclusion of stochastic activity in neural models improves predictions of psychophysical data, in Chapter 6 a model of loudness is derived which can incorporate either the deterministic or the stochastic description of auditory nerve response from Chapter 5. In Chapter 7 the stochastic model of auditory nerve response to electrical stimulation is extended to describe responses to pulse trains. This extended model provides a more precise prediction of the spatiotemporal patterns of response produced by cochlear implant speech processing strategies.

In Chapter 8 the results of these investigations are summarised and drawn together to examine their implications for spatiotemporal coding of sound by cochlear implants.

## Chapter 2

# Frequency coding by cross-fibre interspike intervals

### 2.1 Introduction

It is known that in a single Auditory Nerve (AN) fibre the predominant interval between action potentials is the same as the period of a pure-tone acoustical stimulus for frequencies up to around 600 Hz, and this provides a possible mechanism by which the brain codes frequency (Clark, 1996; Javel and Mott, 1988; Horst et al., 1986; Rose et al., 1967; Wever, 1949). For higher frequencies, refractory effects prohibit a single fibre from firing every cycle, such that the predominant peak in the Interspike Interval (ISI) histogram becomes a multiple of the period of the stimulus. However, the difference between the ISIs in single fibres represents the period of the stimulus for frequencies up to 5 kHz (Rose et al., 1967), although more recent studies indicate that the upper limit is 1.5–3.0 kHz (Clark, Carter, Maffi and Shepherd, 1995; Johnson, 1980). This suggests that ISIs in groups of fibres could be used for coding frequency in this higher frequency range, following the Volley Principle first proposed by Wever (1949). In this chapter interspike intervals across a pair of auditory nerve fibres are investigated to determine how they may code the stimulus frequency.

In a pair of fibres innervating different regions of the cochlea, the distribution of ISIs in each fibre is affected by factors such as the frequency of the pure-tone stimulus, the characteristic frequencies of the two fibres, the firing rate versus stimulus intensity behaviour of the fibres, and the degree of phase-locking in the responses. The ISIs between the fibres is further affected by the phase delay of the travelling wave. It is uncertain how these factors affect frequency coding across fibres.

In this chapter the statistical relationship of spike events between two fibres innervating different regions in the cochlea is investigated using a mathematical model of the cochlea, a hair-

cell/auditory-nerve transduction model and an integral expression for the Cross-Fibre Interspike Interval (CFISI) probability distribution. Given a time-varying acoustical stimulus, the cochlear model simulates the propagating waves in the cochlear fluid resulting in the vibrations of the basilar membrane and the shearing movements of the hair cells. The auditory nerve model then takes the hair cell shearing displacements and converts them into the fluctuating firing probabilities of AN fibres. The cross-fibre interspike interval probability distribution is then calculated from the firing probabilities using an integral expression. The effect of propagation delays on the cross-fibre interspike intervals is studied and its implications for cross-fibre temporal frequency coding are discussed.

## 2.2 The cochlear model

The cochlear model is a linear, active, time-domain model which is based on a frequency-domain model published by Neely and Kim (1986). The cochlea is modelled as a fluid-filled rectangular box with rigid boundaries, and it is separated into upper and lower halves by a flexible cochlear partition. The endocochlear fluid is assumed to be inviscid and incompressible, and the propagation of waves in the fluid is restricted to one-dimension in order to simplify the solution. The cochlear partition is modelled as a 4<sup>th</sup> order spring-mass-and-damper vibration system shown in Figure 2-1. It is driven by the fluid pressure difference across the partition and a frequency dependent active feedback pressure source<sup>1</sup> which is proportional to the hair bundle displacement given by the relative motion of the two masses. A simplified middle ear model is also included to provide the air-to-fluid coupling from the eardrum to the stapes, so that the input stimulus may be specified as the sound pressure at the eardrum or the acceleration of the stapes. The output of the model is the basilar membrane displacement or the hair-cell shearing displacement. In Figure 2-2 the amplitude (top) and phase (bottom) of the hair bundle vibration are plotted as a function of position on the basilar membrane, for a 0 dB SPL pure tone at five different frequencies.

According to Neely (1981), the boundary condition at the stapes may be solved in the time-domain using spatial integration to conserve the volume of fluid in the cochlea. However, this causes stability problems when the waves are reflected from the apex due to the inaccurate model of the helicotrema. To overcome this problem, temporal integration can be used to solve the boundary condition at the stapes, and the helicotrema can either be modelled as zero pressure difference across the apical wall or as a vibrating apical wall with a pure damping impedance.

---

<sup>1</sup>This models the active response of the outer hair cells to the basilar membrane vibration, but does not take into account input to the outer hair cells from efferent fibres originating from the cochlear nucleus.

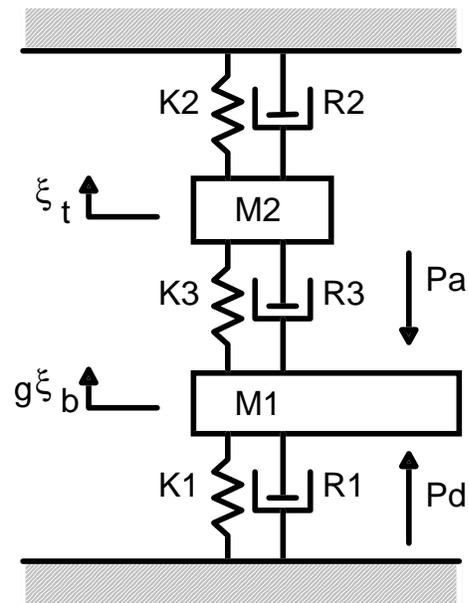


Figure 2-1: Neely and Kim's (1986) lumped component model of cochlear micromechanics. The mass  $M_1$  represents a cross section of the organ of Corti which is attached to rigid bone by stiffness and damping components  $K_1$  and  $R_1$ . The basilar membrane is "driven" by fluid pressure difference  $P_d$  and active pressure source  $P_a$ . The mass  $M_2$  represents a cross section of the tectorial membrane which is attached to rigid bone by  $K_2$  and  $R_2$ . The two masses are coupled by  $K_3$  and  $R_3$ .  $\xi_t$  represents the displacement of the tectorial membrane ( $M_2$ ), and  $g\xi_b$  represents the displacement of the basilar membrane ( $M_1$ ).

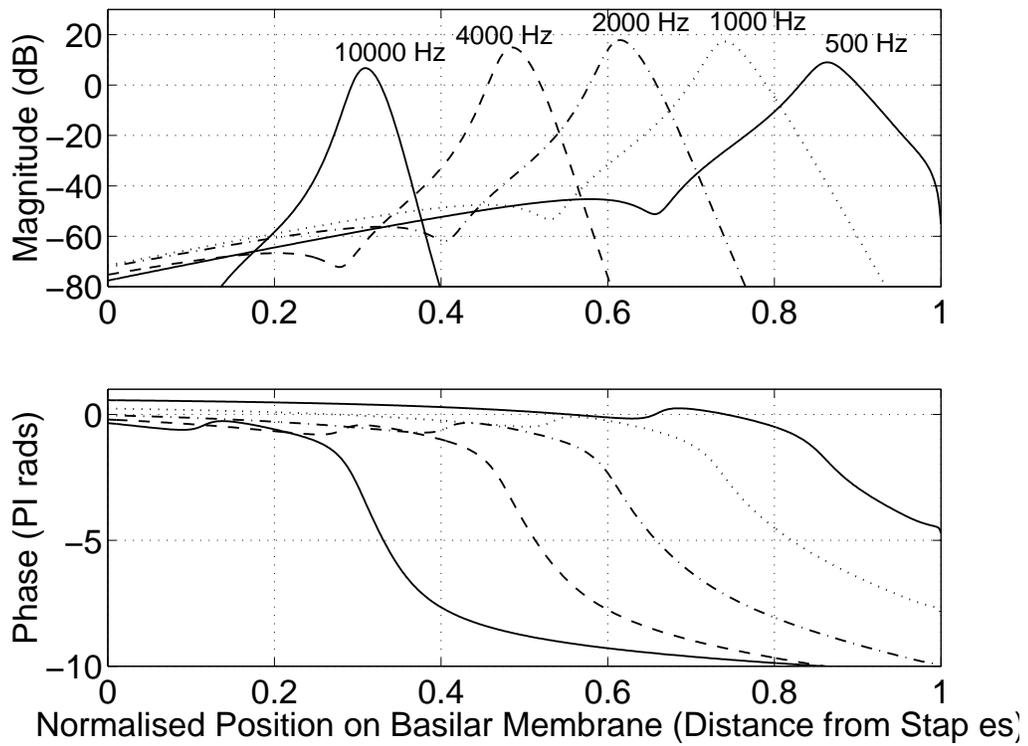


Figure 2-2: Magnitude (top) and phase (bottom) of the hair-bundle vibration versus normalised position on the basilar membrane (distance from stapes), for a 0 dB SPL pure tone at five different frequencies.

## 2.3 The hair-cell/auditory-nerve model

The vibration pattern on the basilar membrane is converted into the fluctuating probability of an action potential being generated using an auditory synapse model based on a model published by Meddis et al. (1990). The model parameters are set to produce an output matching ‘high-spontaneous-rate/low-threshold’ fibres (Meddis et al., 1990). The Meddis et al. model contains three reservoirs of neural transmitters, namely, a transmitter factory, a free transmitter pool and a reprocessing store. The transmitter factory manufactures new neural transmitters and releases them into the free transmitter pool at a rate proportional to the maximum number of transmitters in the free pool. As the basilar membrane vibrates, the instantaneous shearing displacement on the hair cell produces a proportional change in the rate at which the free transmitters are released into the pre-synaptic cleft. Some of the transmitters in the cleft are depleted while others are recovered into the reprocessing store where they are regenerated and released back into the free transmitter pool. The rates of depletion and re-uptake are proportional to the amount of transmitters in the cleft. Finally, the firing probability of the fibre is then obtained by multiplying the amount of transmitters in the cleft by a constant scaling factor.

However, this firing probability does not take into account the refractory period after the fibre has fired. It simply represents the instantaneous probability of an action potential irrespective of when the last action potential occurred. To simulate the effect of an absolute refractory period, the Meddis model generates neural spikes by comparing the instantaneous firing probability with a pseudo-random number, and suppresses all action potentials when the fibre is in the refractory period. With such a model, a post-stimulus time histogram can be obtained by putting the same stimulus through the model repeatedly until a reasonable number of spikes are obtained, which may require a lot of computational time when the probabilities are small. Even more iterations may be required for estimation of the firing probabilities after they have been modified by the refractory period. To overcome this problem, refractory-modified firing probabilities are calculated directly according to Eqn. 22 of Edwards and Wakefield (1990). For a step refractory function with no relative refractory period, the equation can be simplified to:

$$p_k = \begin{cases} s_k & \text{for } k = 1 \\ s_k - s_k \sum_{i=1}^{\min[k-1, m]} p_{k-i} & \text{for } k = 2, 3 \dots n \end{cases} \quad (2.1)$$

where  $p_k$  is the refractory-modified firing probability in the  $k^{th}$  time-bin,  $s_k$  is the unmodified firing probability,  $m$  is the number of bins in the absolute refractory period, and  $n$  is the number of bins in the duration of the stimulus.

## 2.4 Cross-fibre interspike interval probability distribution

Neural firing response can be modelled as a self-exciting point process (Snyder and Miller, 1991), which for many applications is well approximated by an inhomogeneous (non-stationary) Poisson process (Siebert, 1970). The most obvious technique for evaluating the statistical relationship between action potentials in pairs of fibres is to use cross-correlation. However, cross-correlation measures the probability of the interval between all pairs of spikes and not just the interval between successive spikes. An analysis of the information contained in the ISIs of single AN fibres (Siebert, 1970) indicated that efficient estimation of the stimulus frequency based on all intervals in a single fibre, i.e., a complete auto-correlation, would produce *greater* performance than is seen in psychophysical data. Goldstein and Srulovicz (1977) went on to show that the intervals only between *successive* spikes could better explain psychophysical performance. Likewise, in this chapter an expression is developed for the probability distribution of *successive* spikes across a pair of fibres. This is similar to the analysis of discharges recorded simultaneously from pairs of auditory nerve fibres conducted by Johnson and Kiang (1976). However, Johnson and Kiang only performed this analysis on fibres with no phase-locked (synchronised) response, modelled by a *homogenous* (stationary) Poisson process. Our analysis of the model output extends to fibres exhibiting a phase-locked response.

First, define the Poisson rate of fibre  $i$  as  $\lambda_i(t)$ . Then the probability that fibre  $i$  fires between time  $t - d$  and  $t - d + \delta t$  but not again in the subsequent  $d$  seconds is:

$$\lambda_i(t-d)\delta t e^{-\int_{t-d}^t \lambda_i(s) ds} \quad (2.2)$$

Similarly, the probability that fibre  $j$  fires between time  $t$  and  $t + \delta t$  but has not fired in the previous  $d$  seconds is:

$$\lambda_j(t)\delta t e^{-\int_{t-d}^t \lambda_j(s) ds} \quad (2.3)$$

Using (2.2) and (2.3), and assuming statistical independence between the nerve fibres  $i$  and  $j$ , an integral expression is derived which provides the normalised probability of intervals of  $d$  seconds between action potentials of fibre  $i$  and fibre  $j$ :

$$\text{Normalised Pr } [d] = \lim_{T \rightarrow \infty} \frac{\int_d^T \lambda_i(t-d) e^{-\int_{t-d}^t [\lambda_i(s) + \lambda_j(s)] ds} \lambda_j(t) dt}{T-d} \quad (2.4)$$

The auditory models provide the probability of neural firings in small time-bins. When the bins are sufficiently small such that the firing probability of each bin is small, the probability divided by the bin size approximates the Poisson rates of the action potentials. Upon making this approximation, the output is applied to (2.4) to calculate the CFISI probability distribution.

## 2.5 Results

Figure 2-3 shows the CFISI distribution at 500 Hz and 10 kHz for two cases: (i) both fibres at the characteristic place, and (ii) one fibre at the characteristic place and the other at a more apical position. For case (i) at 500 Hz, the first and highest peak is at zero firing delay, meaning there is a high probability for both fibres to fire simultaneously. The fast decay in the distribution indicates that the CFISI is likely to be a small multiple of the period. Also, because of the large variation in the distribution from one peak to the next, there is good phase locking to the positive phase of the sine wave. For case (ii), the fibres are separated and the probabilities are smaller. The position of the first peak is now shifted to the right to a position determined by the phase delay of the travelling wave, but the difference between the peaks remains the same as the period of the stimulus. At 10 kHz, the distribution becomes flatter such that the probabilities for longer intervals are now comparable with shorter intervals. The phase locking between the fibres is also poor as indicated by the shallow troughs between the peaks.

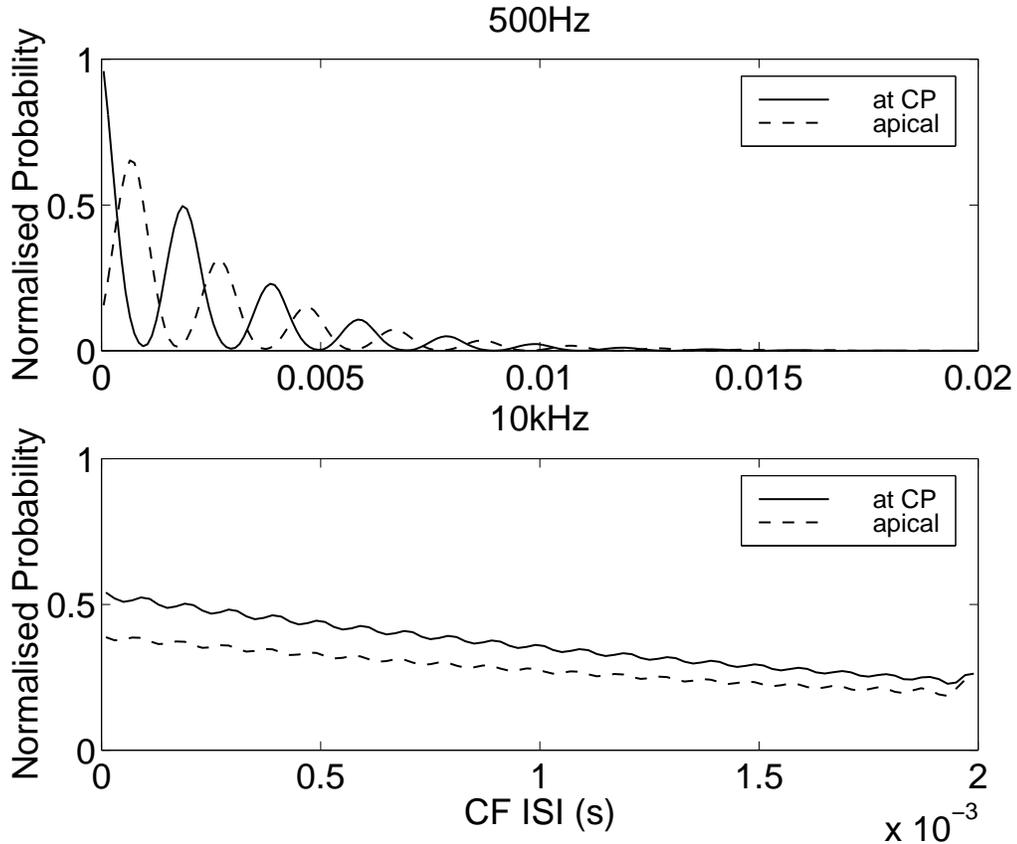


Figure 2-3: Cross-fibre interspike interval probability distribution with both fibres at the characteristic place (solid lines), and with one fibre at the characteristic place and a second at a more apical site (dashed lines), for a 500 Hz (top) and 10 kHz (bottom) pure-tone.

The results plotted in Figure 2-3 indicate that the stimulus frequency could be coded by the dominant cross-fibre interspike interval in a pair of fibres which both originate from the CP, if coincident spikes (the peak corresponding to a cross-fibre interspike interval of zero) are excluded. However, if one of the fibres originates from some other site on the basilar membrane, a shift occurs in the CFISI probability distribution. Consequently, if the dominant non-zero CFISI were to be used to code frequency, this shift would need to be compensated for by some delay line before the CFISI is calculated. However, no delay line would be needed if the two fibres had a spatial separation corresponding to a phase delay which is a multiple of the tone period—in this case the phase of the two responses would be effectively identical and no shift would appear in the CFISI distribution. One restriction on such a coding scheme is the strength of the synchronisation of the CFISI. In order to study the synchronisation between the fibres, the cross-fibre Synchronisation Index (SI) is estimated from the distribution by dividing the fundamental component of the Fourier transform by twice the DC component<sup>2</sup>.

Figure 2-5 shows a plot of the cross-fibre SI against the input frequency with both fibres at the characteristic place. At low frequencies, the cross-fibre SI is high. But for frequencies above 1 kHz, the cross-fibre SI drops rapidly reaching almost zero at 10 kHz. This suggests that temporal frequency coding across fibres may be limited to frequencies below 1–3 kHz. This result is in agreement with 1.5–3.0 kHz upper limit to temporal frequency coding proposed by Clark, Carter, Maffi and Shepherd (1995) and Johnson (1980), rather than the 5 kHz limit of Rose et al. (1967).

Plotted in Figure 2-5 is the cross-fibre SI (where one fibre is at the CP) versus the characteristic frequency of the second fibre, i.e., one fibre is fixed at the characteristic place and the other is moved from basal positions to apical positions. The spread of fibres that are in synchrony is narrower for the high frequencies than for the low frequencies. To investigate the cross-fibre synchronisation in the spatial domain, the cross-fibre SI is plotted against the normalised distance from stapes, as shown in Figure 2-6. It shows that the spread of a pair of fibres that exhibit synchrony in their CFISIs is narrower in the basal region than in the apical region.

A plot of the cross-fibre SI versus the phase difference between a pair of fibres (where one fibre is at the CP) for 5 frequencies from 500 Hz to 10 kHz is shown in Figure 2-7. The results indicate that a pair of fibres with a phase difference smaller than 6–8 radians when the second fibre is on the basal side or 8–10 radians when the second fibre is on the apical side exhibit a good cross-fibre SI for all 5 frequencies. This suggests that only cross-fibre interspike intervals measured from pairs of fibres which are both at the characteristic place (in phase) or which are one cycle ( $2\pi$  radians) out of phase with each other may have sufficient synchronisation and would therefore be suitable for coding the stimulus frequency by the dominant non-zero CFISI.

---

<sup>2</sup>This estimate is exactly equal to the synchronisation index (Johnson, 1980) of a sinusoid fitted to the cross-fibre interspike interval probability distribution.

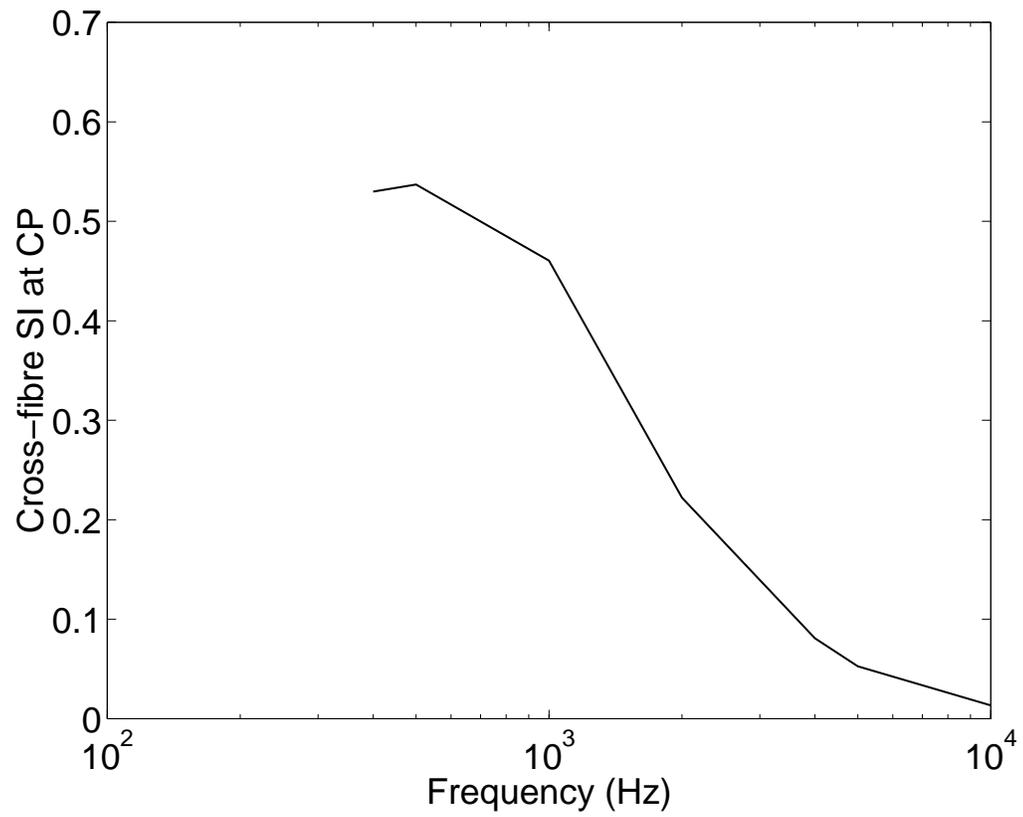


Figure 2-4: Cross-fibre Synchronisation Index (SI) versus characteristic frequency, for two fibres at the characteristic place.

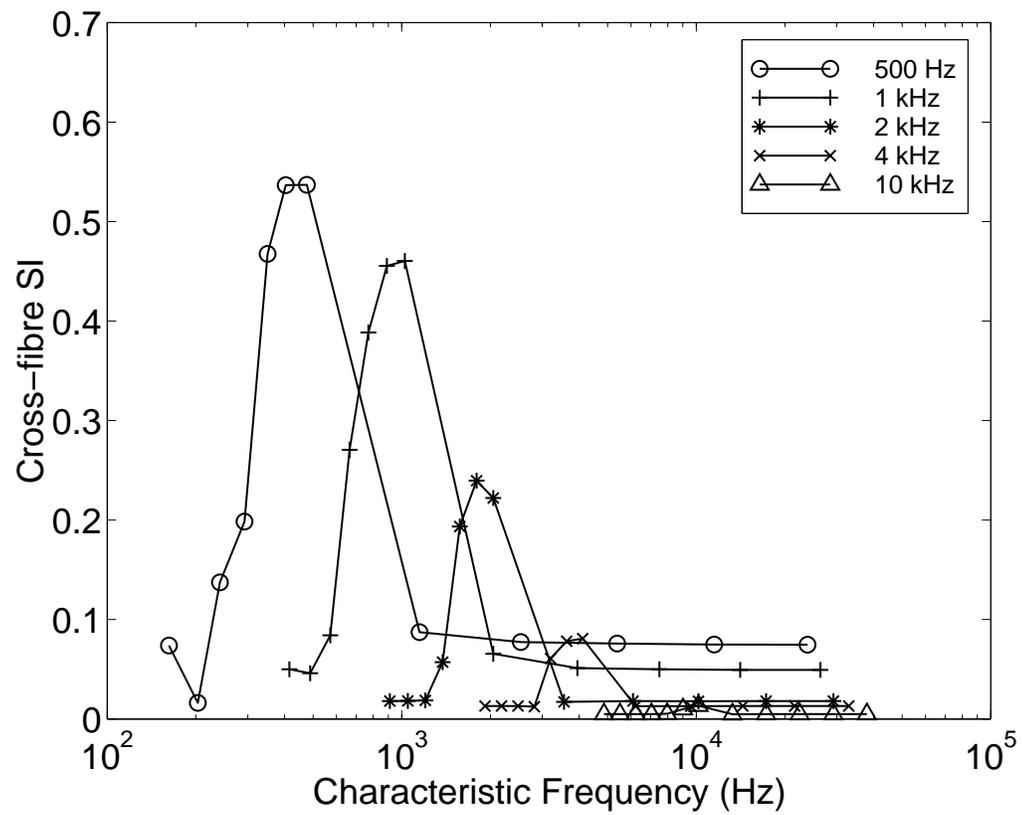


Figure 2-5: Cross-fibre SI (where one fibre is at the CP) versus the characteristic frequency of the second fibre.

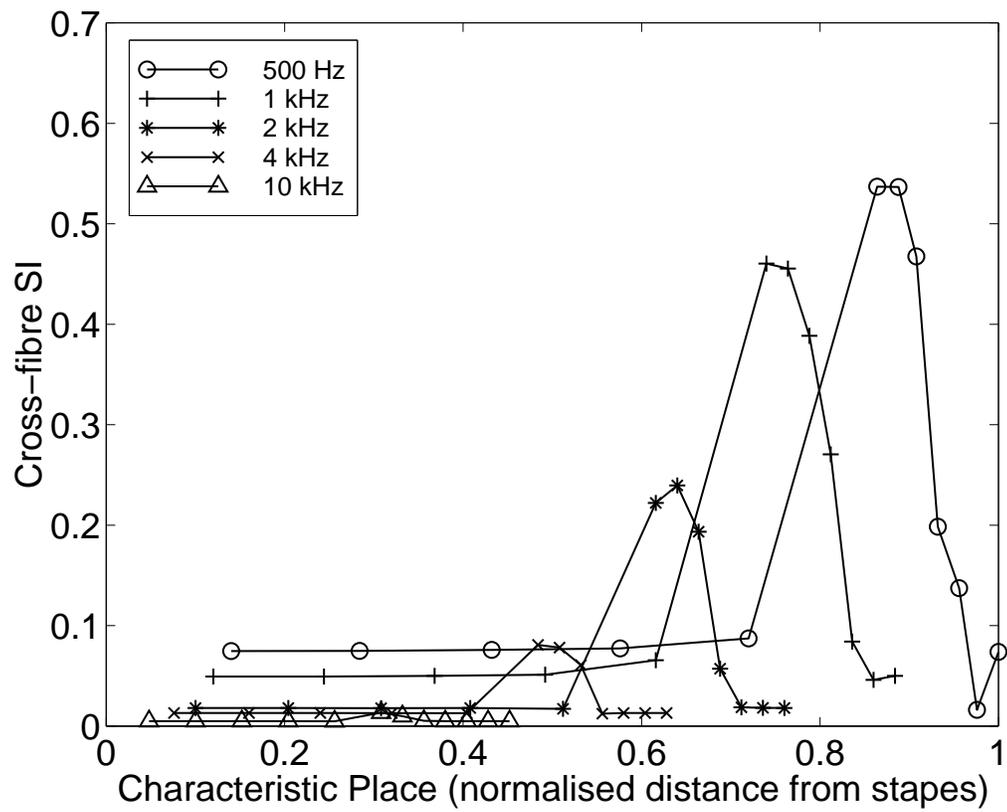


Figure 2-6: Cross-fibre SI (where one fibre is at the CP) versus the characteristic place of the second fibre (normalised distance from stapes).

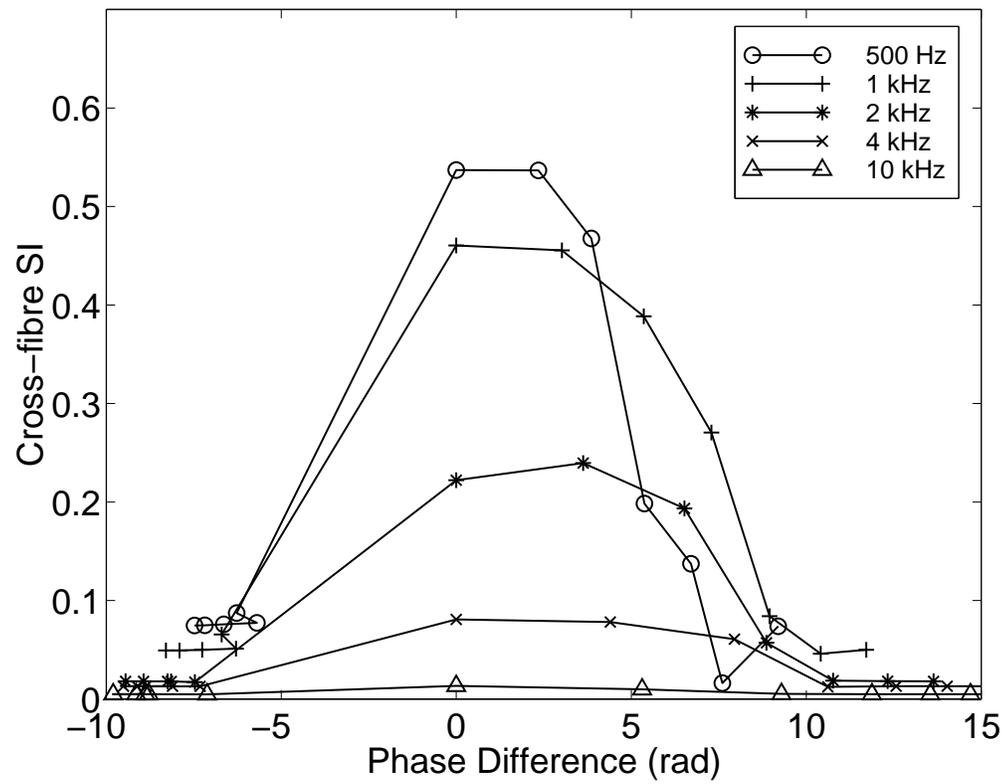


Figure 2-7: Cross-fibre SI (where one fibre is at the CP) versus the phase difference between the two fibres.

## 2.6 Discussion

The results of this chapter show that temporal frequency coding can take place across a pair of fibres in the form of cross-fibre interspike intervals, but it can be affected by factors such as the frequency of the stimulus and the site of origin on the basilar membrane of the two fibres. The cross-fibre synchronisation index suggests that temporal frequency coding may occur across fibres only for frequencies below 1–3 kHz. Furthermore, it suggests that cross-fibre interspike intervals measured from pairs of fibres where both are at the characteristic place (in phase), or where one is at the characteristic place and the other is at a site one cycle ( $2\pi$  radians) out of phase, may have sufficient synchronisation to code the stimulus frequency. Such an arrangement would remove the shift in the cross-fibre interspike interval probability distribution which occurs when the fibres are out of phase by an amount not equal to a multiple of the period. Consequently the stimulus frequency could be coded by the dominant cross-fibre interspike interval, if coincident spikes (a cross-fibre interspike interval of zero) are excluded.

## Chapter 3

# Spatiotemporal summation of auditory nerve firings

### 3.1 Introduction

In the previous chapter combination of activity in pairs of Auditory Nerve (AN) fibres was investigated to determine how it may be used in spatiotemporal coding of a pure tone's frequency. In this chapter the investigation is extended to a spatiotemporal coding mechanism which may combine information from larger populations of fibres, that is, spatiotemporal summation of firings in ensembles of AN fibres (Clark, 1996; Rothman and Young, 1996; Carney, 1994; Rothman et al., 1993; Carney, 1992; Winter and Palmer, 1991; Kim and Parham, 1991; Kim et al., 1990; Viemeister, 1988; Rhode and Smith, 1986; Teich and Lachs, 1979; Rose et al., 1967; Wever, 1949).

The Cochlear Nucleus (CN) is the first section of the brain to process signals coded in the firing patterns of AN fibres. Among the varied cell types and neural connectivities of the CN, it has been found that some groups of AN fibres converge onto individual neurons of the CN and that the response of these neurons appears to be the spatiotemporal summation of the activity of the convergent AN fibres (Paolini et al., 1997; Joris, Carney, Smith and Yin, 1994; Joris, Smith and Yin, 1994; Young et al., 1988; Smith and Rhode, 1987). Previous experimental and modelling work has concentrated on describing particular cell types, but little research has been done to determine which AN fibres provide input to these summing neurons and what effects different spreads of inputs would have on the neurons' outputs. It can be hypothesised that these effects could be significant, because of differences in the response of AN fibres originating from different sites on the Basilar Membrane (BM). In particular, the phase-lag behaviour of the travelling wave produces phase differences in the synchronised (phase-locked) response components of AN fibres, with the greatest changes in phase occurring around the site of maximum excitation.

The purpose of this chapter is to determine theoretically how various degrees of AN fibre convergence onto summing neurons in the CN would affect estimation and discrimination of stimulus features from the summing neurons' outputs. Such an analysis could provide new insights into spatiotemporal summation of AN activity by:

1. determining, given a particular theory of sound coding, what spread of input fibres would be useful (or even optimal) in the processing of AN activity, prompting anatomical and physiological investigation of what degrees of convergence actually exist, and
2. providing hypotheses for the function of different CN cell types, if they are found to have theoretically useful/optimal spreads of inputs.

A variety of approaches have previously been used to model summing neurons (Burkitt and Clark, 1997; Rothman and Young, 1996; Carney, 1994; Rothman et al., 1993; Carney, 1992). In this chapter a parametric model is proposed which permits investigation under various scenarios, such as the number of participating nerve fibres and their response properties, etc. The model is based on a point process model of single-fibre response, coupled with experimentally based parametric descriptions for both the expected rate of firing (the per-stimulus time histogram) and also for the extent of convergence onto the summing neuron. Due to the analytical form of the description, it also facilitates further investigations of information content via theoretical means such as the Cramer–Rao Bound (Kay, 1993) or channel capacity (Guiaşu, 1977).

In Section 3.2, analytical models with the desired generality and parametric descriptions are developed. Based on these models, Section 3.3 details parametric properties of the models and compares an analytical approximation with a more complete numerical model. Section 3.4 investigates how summing neurons with different parameters could be useful in the processing of AN activity, under two key theories of the neural coding of sound, and future directions in experimental and modelling studies are discussed in Section 3.5.

## 3.2 Statistical model of summed AN response

In this section a model is developed to describe the response properties of a neuron which sums the activity of a group of convergent AN fibres. A general (parametric) mathematical description of the response of individual AN fibres is outlined, fitting the parameters via comparison with experimental observations of neural responses. Analytical approximations for the output of the summing neuron are then derived. These approximations permit analysis of the capacity of such structures to convey information about the stimulus, and so form the basis of Sections 3.3 and 3.4.

### 3.2.1 Single AN fibre model

Much progress has been made in the mathematical modelling of AN spike trains. There have been many deterministic models of neural firing, most based on the Hodgkin–Huxley (1952) or Frankenhauser–Huxley (1964) descriptions. These multi-variable, nonlinear differential equation descriptions are able to accurately estimate a fibre’s response to a given stimulus, but due to their highly nonlinear nature have the drawback of being computationally expensive and not easily parametrically analysed (Dean and Lawrence, 1983). Furthermore, it is extremely difficult to incorporate stochastic components of response into such models. However, it is generally accepted that sound features are coded only in the timing and place of the action potentials.

It is therefore assumed that a description of the spike timing in AN fibres will provide sufficient information for the analysis of sound coding. The AN’s response can be closely approximated by a series of stochastically distributed identical spikes, and therefore can be modelled via an unmarked point process model (Snyder and Miller, 1991). Such models only provide information about the statistical properties of the timing of the neural responses and do not calculate the actual membrane potential or ion conductances of the fibre. Thus they are much simpler to analyse and less computationally expensive than the more complete deterministic models.

Point process models range in simplicity from the inhomogeneous Poisson process model through to the doubly-stochastic and self-exciting point process models (Snyder and Miller, 1991). More complex model descriptions often permit greater accuracy of approximation (Teich and Khanna, 1985), but may limit analytical investigation of the model’s behaviour. Within the class of point process models an appropriate model must be chosen based on the behaviour of the physical process and the required degrees of model accuracy and simplicity.

In the case of the AN, the generation of action potentials is affected by recovery (refractory) effects conditioned upon earlier action potentials. This immediately suggests modelling the process via a Self-Exciting Point Process (SEPP) model (Snyder and Miller, 1991). Here the firing probability is related not only to the intensity of the stimulus and firing characteristics of the individual fibre, but also to the time since the fibre last fired (via a recovery function). Although there is limited evidence that certain more general point processes models could be better applied to AN fibre behaviour (Kumar and Johnson, 1993), the SEPP model has been found to form an excellent approximation to neural response (Miller and Mark, 1992), and also permits a reasonable degree of analytical investigation (Irlicht and Clark, 1996). Such models have found much application to the analysis of AN response (Bi, 1989; Miller, 1985; Johnson and Swami, 1983).

Thus, following Irlicht and Clark (1996), Bi (1989), Jones and Tubis (1985), Johnson and Swami (1983) and Gaumond et al. (1982), the following assumption is made:

**Assumption 3.2.1** The stochastic process describing the timing of the action potentials in the  $n^{\text{th}}$  fibre is a SEPP (Snyder and Miller, 1991). Define the number

of spikes (events) to time  $t$  as  $N_t$ , then at any time  $t$ , the time since the last spike equals  $t - t_{N_t}$ . The ‘intensity’ (instantaneous rate) of the point process is equal to  $s_n(t)r_n(t - t_{N_t})$ , where  $s_n(t) \geq 0$  is a stimulus related function, depending on time (determined by the properties of the fibre and of the physical stimulus, i.e., the basilar membrane vibration driving that fibre), and  $r_n(t - t_{N_t}) \geq 0$  is a recovery function, which lowers the rate of action potential generation as a function of  $t - t_{N_t}$ , the time since the last action potential.  $r_n(t)$  is determined by the properties of the fibre and is independent of the size of the stimulus.

**Remark 3.2.1** Analysis of discharges recorded simultaneously from pairs of AN fibres (Johnson and Kiang, 1976) indicates that the responses of AN fibres to a tone and to silence can be described as statistically independent point processes, i.e., the assumption of statistical independence between fibres, which is implicit to Assumption 3.2.1, is an accurate description of the physiological behaviour. Johnson and Kiang’s results imply that, for a healthy cochlea, the initiation of spikes in AN fibres is governed by localised processes specific for each fibre. The relationships between vibrations at different sites on the basilar membrane will cause dependencies and cross-correlations between these localised processes. However, in the SEPP description these are dealt with explicitly by the stimulus related functions for each fibre  $s_n(t)$ . If  $s_n(t)$  is known for each fibre (a method of calculating them is derived in this section), then the resulting stochastic processes can be assumed to be otherwise independent and uncorrelated<sup>1</sup>.

According to Assumption 3.2.1 the rate of neural response is dependent on the specific history of neural spike activity via the recovery function. However, it is possible to define an ‘expected’ rate which is independent of the history of the process. This could be observed as the instantaneous average of the rates of a large number of identical fibres subjected to identical stimuli, or as the Per-Stimulus Time Histogram (PSTH), which is the averaged response of a fibre subjected to repeated identical stimuli.

**Definition 3.2.1**

$$\begin{aligned} \lambda_n(t) &\triangleq E[s_n(t)r_n(t - t_{N_t})|t] \\ &= \text{The PSTH of a fibre with stimulus function } s_n(t) \\ &\quad \text{and refractory function } r_n(t - t_{N_t}) \end{aligned}$$

---

<sup>1</sup>It has been suggested that while this may be the case for a healthy cochlea, certain pathologies of the hair cells may produce further dependencies and cross-correlations in AN activity (Eggermont, 1990). However, this chapter is concerned only with coding of sound in the normal auditory system.

**Lemma 3.2.1** (Jones and Tubis, 1985)

Consider a fibre modelled by a SEPP, and following Assumption 3.2.1, where  $\lambda_n(t)$  is of Definition 3.2.1. Then the expected rate of neural discharge (the PSTH) is given by:

$$\lambda_n(t) = s_n(t) \left[ 1 - \int_0^t \lambda_n(\gamma) [1 - r_n(t - \gamma)] e^{-\int_\gamma^t s_n(\xi) r_n(\xi - \gamma) d\xi} d\gamma \right] \quad (3.1)$$

Lemma 3.2.1 provides a nonlinear integral-equation representation of the expected rate of neural response ( $\lambda_n(t)$ ) in terms of the stimulus and refractory functions. It is not possible to solve (3.1) analytically in the general case, leaving solution to either a series expansion or a special case.

Certain restrictions of either  $r_n(t)$  and/or  $s_n(t)$  cause equation (3.1) to become analytically tractable. Such approximations can provide much information about the neural response in the general case (Irlicht and Clark, 1996; Johnson and Swami, 1983), and also exact representations in special cases. One such approximation is as follows:

**Approximation 3.2.1**  $r_n(t) = U(t - a)$ , where  $U(t)$  is the unit step function, and  $a$  is the non-paralysable dead time (absolute refractory period) of the fibre, i.e., the minimum possible time between neural firings. In this case (3.1) can be expressed as:

$$\lambda_n(t) = s_n(t) \left[ 1 - \int_{\max[0, t-a]}^t \lambda_n(\gamma) d\gamma \right] \quad (3.2)$$

**Remark 3.2.2** The subset of SEPPs which hold to this assumption are called Dead-Time-Modified Poisson Processes (DTMPP). They have found application to diverse areas such as neural systems, particle detection, and optical communications (Prucnal and Teich, 1983; Saleh and Teich, 1982; Vannucci and Teich, 1981; Müller, 1973).

**Remark 3.2.3** A DTMPP with a dead time of zero will cause  $\lambda_n(t)$  of (3.2) to equal  $s_n(t)$ , in which case the process becomes an unmodified Poisson Process.

**Remark 3.2.4** In many cases of interest, Approximation 3.2.1 does not significantly affect the statistics of neural response (Irlicht and Clark, 1996). It is applied here to facilitate calculation of the process variance.

### 3.2.2 Fitting the single AN fibre model to experimental data

Assumption 3.2.1 and associated theory permit the analysis of neural response in terms of the stimulus and refractory functions. However, they do not specify these functions and consequently cannot provide a parametric description for the expected response  $\lambda_n(t)$ . Estimation of  $r_n(t)$  and  $s_n(t)$  functions for a variety of parameters, via minimum description length estimation, can be

used to obtain expressions for the expected rate (Mark and Miller, 1992). However, a prohibitively large number of calculations would need to be performed to obtain estimates for all parametric conditions. An alternative method is to experimentally examine the PSTH of AN fibres subjected to a variety of acoustical stimuli to obtain a general parametric description of  $\lambda_n(t)$ , and apply the SEPP model (with Approximation 3.2.1) to calculate the process variance. This method avoids the time consuming estimation of  $r_n(t)$  and  $s_n(t)$ , by assuming a fixed  $r_n(t)$ , where  $r_n(t) = U(t - a)$ , with ‘ $a$ ’ a constant over all fibres and all stimuli. The result is a computationally inexpensive approximation of neural response, parameterised by factors including the fibre’s origin on the BM and the stimulus frequency and intensity.

### Parameterisation of $\lambda_n(t)$

The discharge rate of a single fibre can be represented by the Fourier series (ten Kate and van Bekkum, 1988; Johnson, 1980):

$$\rho(t) = \bar{\rho} \left( 1 + 2 \sum_{k=1}^{\infty} S_{kf} \sin[2\pi(kft + \phi_{kf})] \right) \quad (3.3)$$

where  $\bar{\rho}$  is the average rate,  $S_{kf}$  is the relative amplitude and  $\phi_{kf}$  the phase of the  $k^{th}$  harmonic of the frequency  $f$ .

Experiments have demonstrated that, for a pure tone, the fundamental frequency component dominates the synchronised response, and therefore it is assumed:  $0 \leq S_{kf} \leq \frac{1}{2}$  for  $k = 1$ , and  $S_{kf} = 0$  for  $k \geq 1$ . Thus the discharge rate of the  $n^{th}$  AN fibre in response to a pure tone can be approximated by a Fourier series with one sine term:

$$\lambda_n(t) = A_n + B_n \sin(\omega t + \phi_n) \quad (3.4)$$

where  $A_n (= \bar{\rho})$  is the average firing rate,  $B_n (= 2\bar{\rho}S_f)$  is the amplitude of the fundamental frequency component,  $\omega (= 2\pi f)$  is the tone’s cyclic frequency and  $\phi_n$  is the phase delay in the firing-time distribution of the  $n^{th}$  fibre.

For any fibre, the Synchronisation Index (SI) measures the amount with which neural response is synchronised to the phase of the stimulus. It can be experimentally determined as explained in (Johnson, 1980; Anderson, 1973), and for the form of neural responses described by (3.4) is equal to:

$$SI = \frac{B_n}{2A_n} \quad (3.5)$$

**Remark 3.2.5** If the SI of a fibre is significantly higher than one-half, or if ‘peak-splitting’ (the presence of a second harmonic component (Johnson, 1980)) occurs, then

(3.4) will not sufficiently model the behaviour of that particular fibre, and a more complex Fourier series as expressed in (3.3) must be used. Although such expressions have been developed (Anderson, 1973), these are still periodic in nature and will produce similar constructive or destructive superposition in a summing neuron to that seen using this simple model. This is indicated by the computer model in Section 3.3.1.

Equation 3.4 is a general description of neural response, including a parameterisation of average rate, synchronisation, and phase delay. Together with the SEPP theory of Section 3.2.1 it forms the basis of the neural models of this chapter. Experimental results and computer models can be examined to determine the parameters (here  $A_n$ ,  $B_n$  and  $\phi_n$ ).

The relationship between the single-fibre response and the signal intensity and frequency has been investigated experimentally in (Johnson, 1980), producing Observations 1–4. Observations 5 and 6 arise by fitting Equation (3.4) to the output of the computer model described in the previous chapter. The transfer function of the basilar membrane section of this model, plotted in Figure 3-1, displays the cochlea's filter bank and phase lag behaviour.

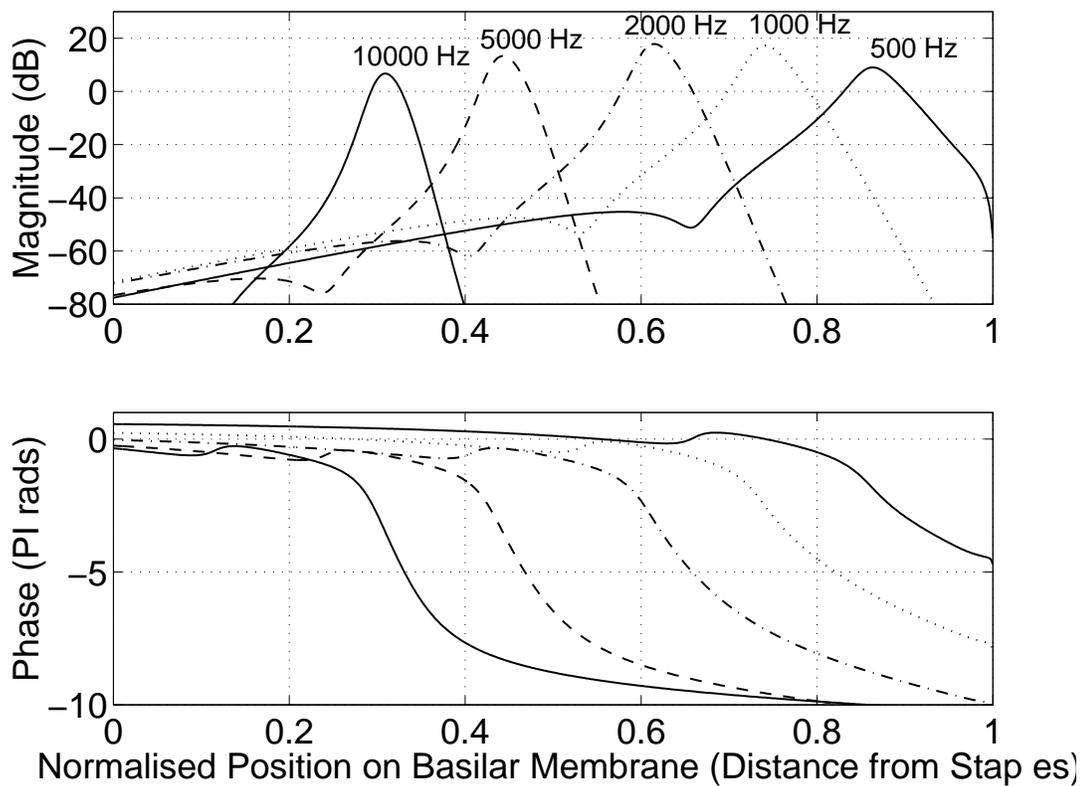


Figure 3-1: Basilar membrane vibration amplitude (top) and phase lag (bottom) curves for five frequencies between 500 Hz and 10 kHz.

### Experimental observations of $\lambda(t)$

1. Both  $A_n$  and  $B_n$  increase fairly monotonically with signal intensity, but the SI appears to begin decreasing again at higher levels ( $\geq 50$  dB SPL), due either to peak-splitting or to a skewed waveform.
2.  $B_n$  appears to decrease monotonically with Characteristic Frequency (CF), from its maximum value of  $A_n$  at low CFs ( $\leq 1$  kHz), to a value of 0 at higher CFs ( $\geq 5$  kHz).
3. The average rate,  $A_n$ , also displays a ‘long-term’ adaptation effect, where it decreases quite rapidly in the first few seconds of the presentation of a constant tone, with a gradual decline over the next few minutes to a relatively stable value. The SI does not display this behaviour, however, remaining approximately constant over the entire presentation time. This indicates that  $B_n$  also undergoes adaptation, decreasing at a rate which maintains its proportionality to  $A_n$ .
4. Units with low spontaneous discharge rates tend to have a higher maximum SI than units with high spontaneous discharge rates.
5.  $A_n$  and  $B_n$  are at their maximum at CF, and drop off as the frequency departs from CF. This frequency selectivity is normally displayed by means of tuning curves, and can also be displayed by plotting  $A_n$  and  $B_n$  as functions of stimulus frequency for a particular sound intensity. This is shown in Figure 3-2 for a model of identical ‘high-spontaneous-rate/low-threshold’ auditory nerve fibres at different points along the basilar membrane.
6. The phase lag in the basilar membrane vibration as shown in Figure 3-1 produces systematic differences in the phase of the synchronised response in each fibre ( $\phi_n$ ).

If it can be assumed that a summing neuron receives input from identical fibres, this will greatly simplify the derivation of a model of that neuron. Auditory nerve fibres do differ in their spontaneous response, and also in their onset and saturation thresholds. However, experiments have shown that fibres with low spontaneous rates and high thresholds show enhanced responses to the envelopes of complex sounds, while fibres with higher spontaneous rates and low thresholds respond to the temporal fine structure (Langner, 1992; Horst et al., 1986). Because AN fibres with similar input/output properties appear to convey similar temporal information about sound features, the summing-neuron model is restricted to the case where it receives input from identical AN fibres.

**Assumption 3.2.2** It is assumed that all AN fibres converging on to any one summing neuron are identical, except in the site on the basilar membrane from where they originate.

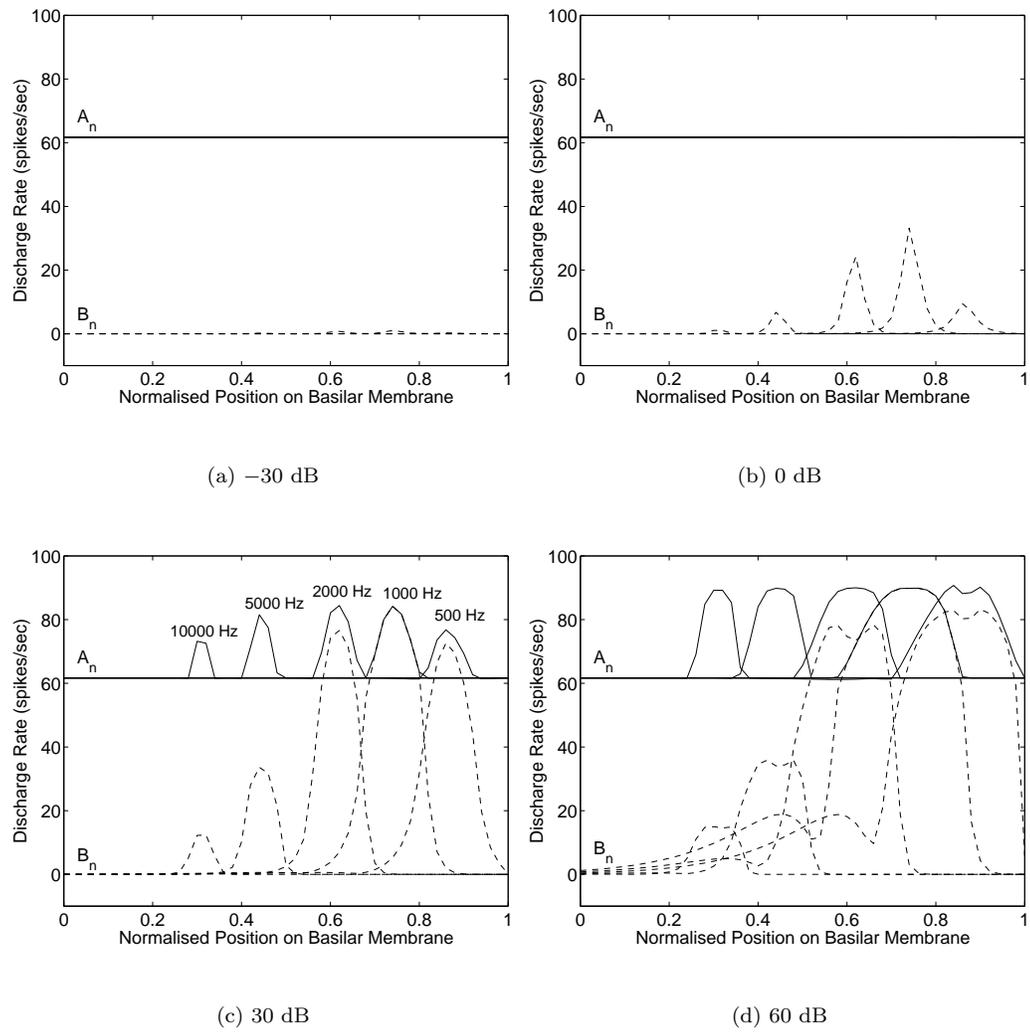


Figure 3-2:  $A_n$  (solid lines) and  $B_n$  (dashed lines) versus normalised position on the basilar membrane for five different stimulus frequencies, as labelled in pane (c), at sound levels of -30 dB (a) to 60 dB (d).

Based on Assumption 3.2.2 and Observations 5 and 6, two further approximations are proposed:

**Approximation 3.2.2** Define the Characteristic Place (CP) as the position on the basilar membrane corresponding to the CF. Then,  $A_n$  and  $B_n$  are approximately constant for identical fibres close to the CP and will therefore be denoted simply as  $A$  and  $B$ .

**Approximation 3.2.3** The phase delay  $\phi_n$  changes linearly with slope  $\alpha$  over small distances around the CP. For  $N$  identical fibres spread out evenly over distance  $d$  on the BM, the phase delay of the  $n^{\text{th}}$  fibre can be expressed by:  $\phi_n = \alpha d \frac{n}{N}$ , and therefore (3.4) can be approximated by:

$$\lambda_n(t) = A + B \sin\left(\omega t + \alpha d \frac{n}{N}\right) \quad (3.6)$$

**Remark 3.2.6** The approximation of  $B_n$  close to the CP by a constant improves with increasing sound intensity (see Figure 3-2).

**Remark 3.2.7** Sensitivity of the model to Approximations 3.2.2 and 3.2.3 can be examined via application of the computer model described in the previous chapter which does not require these approximations. Comparisons of results generated from the analytical model and from the computer model are given in Section 3.3.1.

### 3.2.3 Calculation of the variance in single fibre response

The variance in single fibre response over an arbitrary time period  $\tau$  can be calculated in the following ways:

1. It can be estimated for any stimulus function by computer simulation of the SEPP process. However, this is computationally expensive and does not permit easy parametric analysis.
2. It can be calculated exactly via analytical expressions for the case of a constant stimulus function. This occurs when both the incident tone is at a constant intensity, and the SI of a single fibre is zero. The synchronised response of a single fibre is at a maximum at low frequencies and drops rapidly when the frequency rises above 1 kHz, to a constant firing rate for frequencies above approximately 5 kHz.
3. It can be approximated analytically for the case of a sinusoidally modulated  $s_n(t)$  where the stimulus function varies slowly relative to the dead time (Vannucci and Teich, 1978). These conditions will be satisfied for low-CF fibres (CF  $\ll$  1 kHz) at low sound intensities (see Figure 3-2) and for mid-CF fibres (1 kHz  $<$  CF  $<$  5 kHz).

**Lemma 3.2.2** Consider a single fibre modelled by a DTMPP (Assumption 3.2.1 and Approximation 3.2.1) which has reached an equilibrium state (Cox, 1962). Define  $s_n(t)$  according to Assumption 3.2.1 and  $\lambda_n(t)$  according to Definition 3.2.1. Define  $N_n(t, \tau)$  as the number of spikes in the  $n^{\text{th}}$  fibre counted in the time interval  $[t - \tau, t]$  and ‘ $a$ ’ as the dead time.

In the case where  $s_n(t) = s_0$ , a constant:

$$\begin{aligned} \text{var}[N_n(t, \tau)] &= \frac{2s_0}{1 + as_0} \sum_{k=1}^{\infty} \frac{s_0^k}{(k-1)!} U(\tau - ak) \int_0^{\tau - ak} [\tau - ak - \gamma] \gamma^{k-1} e^{-s_0 \gamma} d\gamma \\ &+ \frac{\tau s_0}{1 + as_0} - \left( \frac{\tau s_0}{1 + as_0} \right)^2 \end{aligned} \quad (3.7)$$

$$\begin{aligned} \lim_{\tau \rightarrow \infty} \text{var}[N_n(t, \tau)] &= \frac{s_0}{(1 + as_0)^3} \tau \\ &= \lambda_n(t) (1 - a\lambda_n(t))^2 \tau \end{aligned} \quad (3.8)$$

In the case where  $s_n(t)$  varies slowly relative to the dead time (Vannucci and Teich, 1978):

$$\begin{aligned} \lim_{\tau \rightarrow \infty} \text{var}[N_n(t, \tau)] &= \int_{t-\tau}^t \frac{s_n(u)}{(1 + as_n(u))^3} du \\ &= \int_{t-\tau}^t \lambda_n(u) (1 - a\lambda_n(u))^2 du \end{aligned} \quad (3.9)$$

**Proof:** see Appendix A.

**Remark 3.2.8** Numerical analysis indicates the conditions under which the asymptotic variance (3.8) estimates the variance (3.7) to within 10%. This accuracy is obtained in the following situations:

- for all values of  $\tau$ , when  $s_0 = 10$  spikes/second,
- for  $\tau > 1$  ms, when  $s_0 = 100$  spikes/second,
- and for  $\tau > 8$  ms, when  $s_0 = 1\,000$  spikes/second.

**Approximation 3.2.4** Just as  $\lambda_n(t)$  describes the expected rate of single fibre response at time  $t$ ,  $\sigma_n^2(t)$  is defined as the variance in the single fibre rate at time  $t$ . Following Lemma 3.2.2, the variance is approximated via the expression:

$$\sigma_n^2(t) \approx \lambda_n(t) (1 - a\lambda_n(t))^2 \quad (3.10)$$

**Remark 3.2.9** From Lemma 3.2.2, (3.10) is a general approximation for any slowly varying  $\lambda_n(t)$ . (3.6) may be substituted for  $\lambda_n(t)$  in (3.10) to obtain a specific expression of the variance for that case. This expression cannot be further simplified, so it is left in the form of (3.10).

### 3.2.4 A summing-neuron model

In the Introduction the biological evidence was discussed for summing neurons in the CN, which could be used in the spatiotemporal summation of AN responses as proposed by a number of theories of sound coding. A model is now developed of such a summing neuron, which receives input from identical convergent AN fibres described by the single-fibre model developed in the previous section. Expressions for the mean and variance of the model's output are developed as functions of stimulus parameters and degree of AN fibre convergence.

#### Spatial summation of AN activity

Following Assumption 3.2.1, single-fibre responses are described by *statistically independent* stochastic processes<sup>2</sup>, and consequently the mean and variance of the summed response can easily be calculated.

Define  $S(t)$  as the sum of the rates of  $N$  fibres spread out symmetrically around the CP over a distance  $d$  on the BM, and  $W_n$  as the weighting of the  $n^{\text{th}}$  fibre in the sum. Then under Assumption 3.2.2 and Approximations 3.2.1, 3.2.2, 3.2.3 and 3.2.4, the mean of  $S(t)$  can be expressed as:

$$\mu_S(t) = \sum_{n=0}^{N-1} W_n \left[ A + B \sin \left( \omega t + \alpha d \frac{n}{N} \right) \right] \quad (3.11)$$

and its variance as:

$$\sigma_S^2(t, \cdot) \approx \sum_{n=0}^{N-1} W_n \lambda_n(t) (1 - a \lambda_n(t))^2 \quad (3.12)$$

**Approximation 3.2.5** Input fibres are given equal weighting in the summing-neuron model.

**Remark 3.2.10** It is likely that many other distributions of weighting of neural response are present in summing neurons. Any such distribution could easily be incorporated into the results which follow via an appropriate weighting function.

---

<sup>2</sup>See Remark 3.2.1 for a discussion of the conditions under which statistical independence can be assumed.

### Integral approximation of summed response

In order to permit straightforward analysis of the parametric dependence of  $S(t)$ , an integral approximation of  $S(t)$ , defined as  $I(t)$ , is introduced. From (3.6), the rate per unit distance, where the nerve fibre density is  $\frac{N}{d}$ , can be expressed as a continuous function of time,  $t$ , and of position on the BM,  $x$ :

$$\lambda(x, t) = \{A + B \sin(\omega t + \alpha x)\} \frac{N}{d} \quad (3.13)$$

**Approximation 3.2.6**  $S(t)$  can be approximated by  $I(t)$  via integration of the rate as described in (3.13) over distance  $d$ .

$$\begin{aligned} \mu_I(t) &= \int_0^d \{A + B \sin(\omega t + \alpha x)\} \frac{N}{d} dx \\ &= NA + NB \operatorname{sinc}\left(\frac{\alpha d}{2\pi}\right) \sin\left(\omega t + \frac{\alpha d}{2}\right) \end{aligned} \quad (3.14)$$

and

$$\sigma_I^2(t) \approx \int_0^d \lambda(x, t) (1 - a\lambda(x, t))^2 dx$$

where  $\operatorname{sinc}(x) \triangleq \frac{\sin(\pi x)}{\pi x}$

### Conditions for approximation of $\mu_S(t)$ by $\mu_I(t)$

In Teich and Lachs (1979) it was assumed that the fibres and the distribution of characteristic frequencies are so dense that it is reasonable to consider the characteristic frequencies as a continuum. However, if a number of coding mechanisms exist in parallel, then it is possible that not all the fibres in a particular region contribute to each mechanism. In this case, (3.14) cannot be used unless  $\mu_I(t)$  is a reasonable approximation to  $\mu_S(t)$ . To this end, a bound for the difference between the integral and the sum is determined.

**Claim 3.2.1** The relative difference (normalised by the long-term average of  $\mu_I(t)$ ) between the integral and the sum is bound by:

$$\rho \triangleq \frac{|\mu_I(t) - \mu_S(t)|}{\bar{\mu}_I} \leq \left| \frac{\alpha B d}{2NA} \right| \quad (3.15)$$

**Proof:** see Appendix A.

**Remark 3.2.11** The largest possible error will occur when  $B = A$ , and when  $\alpha$  is at a maximum. From the computer model and physiological data,  $B \leq A$  and,

$\alpha \leq \frac{-2\pi}{0.15}$  rad/cm. Therefore an upper bound for the relative error of (3.15) is:  $\rho \leq \frac{d\pi}{0.15N}$ , where  $d$  is measured in centimetres.

Thus, it is possible to guarantee a maximum error of 5% when  $\frac{N}{d} \geq 418.9\text{cm}^{-1}$ , or 1% when  $\frac{N}{d} \geq 2\,094.4\text{cm}^{-1}$ .

**Remark 3.2.12** In the human cochlea there are approximately 30 000 nerve fibres innervating the BM which has a length of about 3.5 cm. This means that the average fibre density is  $8\,570\text{ cm}^{-1}$ . Even if only a quarter of the nerve fibres within any given distance,  $d$ , converge onto a summing neuron, the error will still be less than 1 %. Therefore, the integral expression  $\mu_I(t)$  can be used as an estimator of the sum  $\mu_S(t)$ , except where  $\frac{N}{d}$  is less than  $418.9\text{cm}^{-1}$ .

**Remark 3.2.13** There exist two cases where  $\mu_I(t)$  is exactly equal to  $\mu_S(t)$ , occurring when  $\alpha d = -2\pi$ , or when  $d = 0$ . These are examined in detail in Sections 3.4.1 and 3.4.2.

### Temporal integration

Summing neurons have classically been described by ‘integrate-and-fire’ models (Burkitt and Clark, 1997), where a discharge on any input fibre produces an Excitatory Post-Synaptic Potential (EPSP) in the summing neuron. These EPSPs are then integrated over time by a ‘perfect integrator’ (in which the decay of the potential across the membrane is ignored), by a ‘shot-noise’ model (Carney, 1994; van Stokkum and Gielen, 1989) (in which each EPSP is represented by an instantaneous increase in postsynaptic voltage that decays exponentially over time), or by an integrator which includes a more physiologically accurate description of the EPSP waveform. An analytical description of such temporal integration can easily be obtained for the summed AN response by using the perfect integrator approximation of the integrate-and-fire behaviour, with a rectangular integration window of duration  $\tau$ . This produces a stochastic counting process (Snyder and Miller, 1991).

Defining  $R(t, \tau)$  as the integral of  $I(t)$  over the period  $[t - \tau, t]$ , or  $R(t, \tau) \triangleq \int_{t-\tau}^t I(u) du$ :

$$\mu_R(t, \tau) = NA\tau + NB\tau \text{sinc}\left(\frac{\alpha d}{2\pi}\right) \text{sinc}\left(\frac{\omega\tau}{2\pi}\right) \sin\left(\omega t + \frac{\alpha d}{2} - \frac{\omega\tau}{2}\right) \quad (3.16)$$

and

$$\sigma_R^2(t, \tau) \approx \int_{t-\tau}^t \int_0^d \lambda(x, u) (1 - a\lambda(x, u))^2 dx du$$

**Remark 3.2.14** The temporal integration window duration  $\tau$  will be dependent primarily on the membrane time-constant, which for ventral CN cells is on the order

of 0.5–10 ms (Manis and Marx, 1991), but will also be affected by the EPSP rise-time and dendritic extent. The summing model with a relatively small integration time corresponds to what is classically termed a coincidence detector (Carney, 1994; van Stokkum and Gielen, 1989).

**Remark 3.2.15** The model developed thus far gives a simple yet powerful analytical description of spatiotemporal summation of AN activity, which will permit investigation of the effects of AN fibre convergence on the output of a summing neuron. Integrate-and-fire models typically incorporate some form of nonlinear input/output mechanism (thresholding) in the generation of output spikes, but inclusion of such a thresholding mechanism will greatly complicate the analytical description and may confound the investigation of the effects of the other model parameters. Therefore, in investigating the model’s response properties (Section 3.3), only general theoretical considerations of what effects a thresholding mechanism would have on the model’s output are included. Of particular interest are the results of (Burkitt and Clark, 1997), which suggest that it is the nonlinear input/output (thresholding) mechanism which produces the enhancement of synchrony found in other models (Rothman and Young, 1996; Rothman et al., 1993) and in physiological data (Joris, Carney, Smith and Yin, 1994; Joris, Smith and Yin, 1994)—such effects are considered in the following sections where appropriate.

### 3.3 Response properties of the summing-neuron model

This section examines the response properties of the model developed in the previous section. Firstly, the amount of synchronisation in the summing neuron’s output is calculated as a function of the stimulus and neural parameters. Secondly, the ability to detect or estimate acoustical features from the output of a summing neuron is investigated. This is then related to commonly used statistical and psychophysical measures. These provide important measures of the model’s parametric behaviour which are used in Section 3.4 to investigate how summing neurons with different parameters could be useful in the processing of AN activity, according to key theories of the neural coding of sound.

#### 3.3.1 Synchronisation in the model output

The synchronised (phase-locked) response component of single AN fibres is a significant feature of AN firing patterns. Consequently, many theories have been developed which propose various uses for this synchronisation in the coding of sound features. The Volley Principle, first proposed by Wever (1949), is one such theory which utilises both the synchronised response of AN fibres

and the spatial summation of population responses. Furthermore, physiological studies (Joris, Carney, Smith and Yin, 1994; Joris, Smith and Yin, 1994) have found CN cell types which appear to enhance synchronisation through spatiotemporal summation of AN firings. To investigate the effects of model parameters on synchronisation in the model's output, a measure of synchronisation in the summed firings of a population of fibres is introduced, the Population Synchronisation Index (PSI).

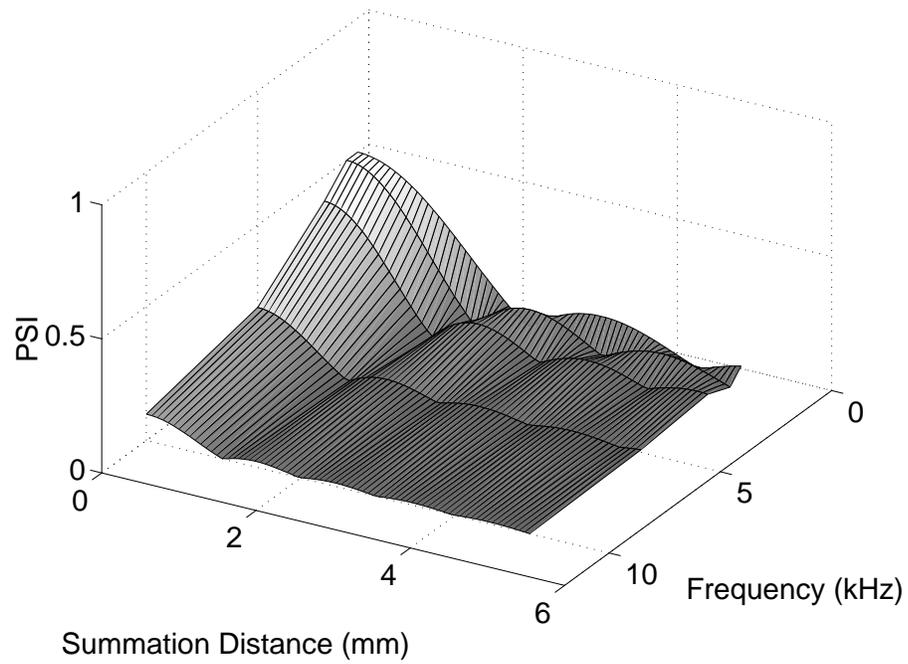
The integral approximation of the mean summed firings of a population  $\mu_R(t, \tau)$ , as described by (3.16), has the same form as the expression for the expected single-fibre discharge rate, given by (3.4). Therefore an expression for the synchronisation of the summed response, the PSI, can be derived which is directly analogous to the single fibre SI as described by (3.5):

$$\text{PSI} = \left| \frac{B}{2A} \text{sinc} \left( \frac{\alpha d}{2\pi} \right) \text{sinc} \left( \frac{\omega \tau}{2\pi} \right) \right| \quad (3.17)$$

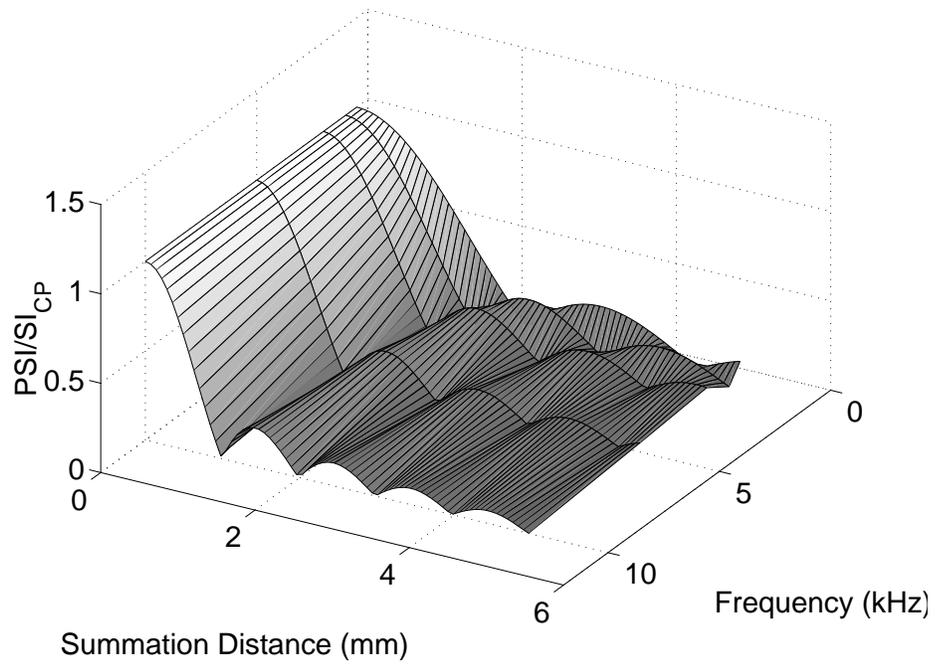
In Figure 3.3(a), the PSI is plotted as a function of summation distance and tone frequency, for a spatiotemporal summing neuron centred at the CP and with  $\tau \rightarrow 0$ . The input to this neuron is the response of identical 'high-spontaneous-rate/low-threshold' AN fibres as shown in Figure 3-2, for a stimulus intensity of 60 dB. The maximum value the PSI can take is  $\frac{B}{2A}$ , which is equal to the SI of fibres at the CP. Because the sinc function has a maximum value of 1, input from fibres away from the CP, measured by the spread of inputs  $d$ , and/or temporal integration of responses, measured by  $\tau$ , can only decrease the synchronisation. This is illustrated in Figure 3.3(b), where the PSI normalised by the SI of single fibres at the CP.

**Remark 3.3.1** The results of Burkitt and Clark (1997) indicate that a nonlinear transformation (thresholding) of the summed response could recover some of the lost synchronisation. Furthermore, if the spread of fibres is restricted to a very small region around the CP and the temporal integration time constant is also very small, then the total effect could be an overall enhancement in the synchronisation at the output.

The analytical model relies on approximation of  $\lambda_n(t)$  by a Fourier series with one sine term (Equation (3.4)) and on approximation of the values of  $A_n$ ,  $B_n$  and  $\phi_n$  for a local population of fibres (Approximations 3.2.2 and 3.2.3). Using the computer model described in the previous chapter, the waveform of  $\lambda_n(t)$  is a better estimate of the expected discharge rate than (3.4), and the waveform is calculated explicitly for each fibre. Thus the computer model can be used to verify the sensitivity of results obtained from the analytical model under these approximations. In Figure 3-4, both the absolute and the normalised PSI of Figure 3-3 are predicted directly by the output of the computer model with the spatiotemporal sum calculated numerically, where  $\tau \rightarrow 0$ . The features which are investigated in Section 3.4 occur in both figures, consequently justifying the approximations for the stated task.

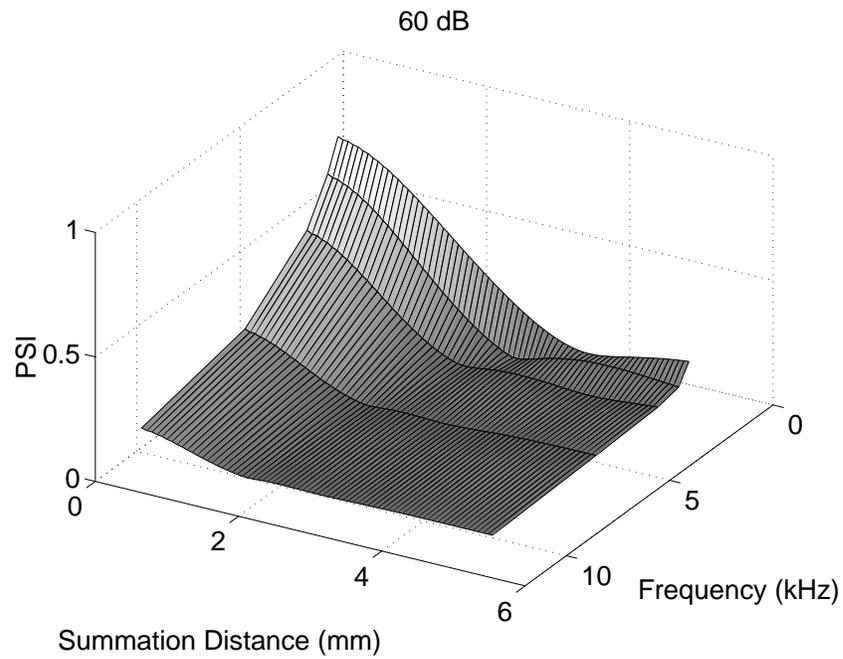


(a) PSI

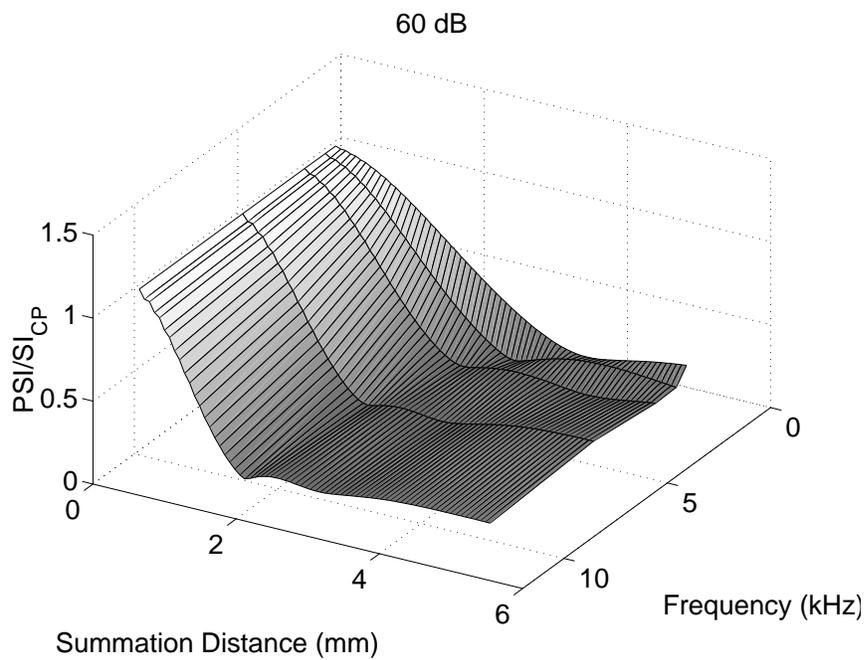


(b) PSI normalised by the SI of single fibres at the CP

Figure 3-3: Analytical results: (a) PSI and (b)  $PSI/SI_{CP}$  versus summation distance and tone frequency, for a spatiotemporal summing neuron centred at the CP and with  $\tau \rightarrow 0$ .



(a) PSI



(b) PSI normalised by the SI of single fibres at the CP

Figure 3-4: Computer simulation results: (a) PSI and (b)  $\text{PSI}/\text{SI}_{\text{CP}}$  versus summation distance and tone frequency, for a spatiotemporal summing neuron centred at the CP and with  $\tau \rightarrow 0$ .

The features are as follows:

1. For an integration distance close to zero, the output of the summing neuron is highly synchronised at low frequencies, and this synchronisation decreases monotonically with increasing frequency.
2. The PSI tends to decrease with increasing integration distance and exhibits oscillations over the integration distance producing local minima and maxima.

### 3.3.2 Mean-to-standard-deviation ratio

The summing-neuron model permits detection or estimation of the properties of a given sound as functions of the neuron's expected (instantaneous mean) response. However, because of the stochastic nature of neural firing, the brain does not obtain input in the form of the expected response, but rather receives a single realisation of the process. The distribution of such realisations determines the extent to which the expected response, and consequently the sound properties, can be estimated. One method of quantifying this is via analysis of the magnitude of the variance (or standard deviation) relative to the mean.

The mean-to-variance ratio (MVR) has previously been used as a measure of the detectability of a counting process similar to the summing-neuron model (Prucnal and Teich, 1983). However, the MVR is not a true unitless ratio. A better alternative is the mean-squared-to-variance ratio (MSVR) (Miller et al., 1987), or its square-root, the mean-to-standard-deviation ratio (MSDR), because they are true unitless ratios. The MSDR is utilised in this chapter. Variants of the MSDR such as confidence intervals and detectability measures are also used extensively in a number of disciplines. A large MSDR implies a good detectability or discrimination of a particular signal, or a tight confidence interval.

#### Application of the MSDR to calculation of confidence intervals

A Confidence Interval (CI) is the probability that a random variable (in this case the actual number of spikes counted by the summing-neuron) lies within a certain interval centred around the mean. The Tchebycheff Inequality (Papoulis, 1991) uses the MSDR to provide a bound on the process CI where the distribution is unknown.

**Lemma 3.3.1** Consider a process with mean  $\mu_R(t, \tau)$  and variance  $\sigma_R^2(t, \tau)$ . Define  $N_R(t, \tau)$  as the number of spikes counted in the period  $[t - \tau, t]$  and  $(\mu_R(t, \tau) - \beta\mu_R(t, \tau), \mu_R(t, \tau) + \beta\mu_R(t, \tau))$  as the interval estimator of  $N_R(t, \tau)$ , then:

$$\Pr \{ \mu_R(t, \tau) - \beta\mu_R(t, \tau) \leq N_R(t, \tau) \leq \mu_R(t, \tau) + \beta\mu_R(t, \tau) \} \geq 1 - \left( \frac{\sigma_R(t, \tau)}{\beta\mu_R(t, \tau)} \right)^2 \quad (3.18)$$

**Proof:** Follows immediately from the Tchebycheff Inequality (See (5-57) of Papoulis (1991) for instance).

**Remark 3.3.2** The relative error  $\beta$ , and consequently the interval estimator  $(\mu_R(t, \tau) - \beta\mu_R(t, \tau), \mu_R(t, \tau) + \beta\mu_R(t, \tau))$ , can be calculated by setting the right hand side of (3.18) to the minimum CI desired. Plotted in Figure 3-5 are expected counts (solid lines) and confidence intervals of at least 75% (dotted lines) for the number of spikes counted in 5 ms by a summing neuron with 10 input fibres, for spreads of inputs creating a zero PSI (top) and a non-zero PSI (bottom).

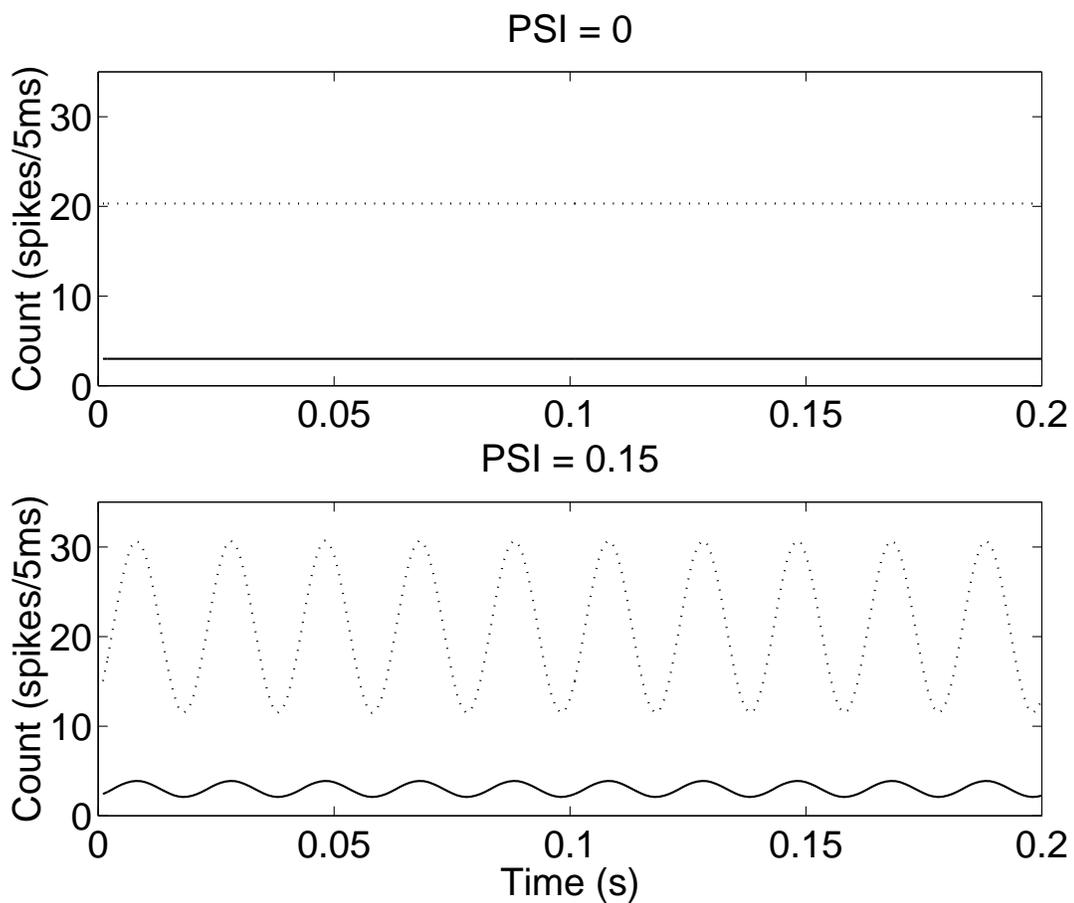


Figure 3-5: Expected counts (solid lines) and confidence intervals of at least 75% (dotted lines) for the number of spikes counted in 5 ms by a summing neuron with 10 input fibres. Top: Zero PSI. Bottom: Non-zero PSI.

**Remark 3.3.3** In the case of a summing neuron which receives input from 30 or more AN fibres, the distribution of the counting process can be approximated by a normal distribution via the central limit theorem. This permits the calculation of CIs tighter than those obtainable from the Tchebycheff Inequality.

### Application of the MSDR to calculation of psychophysical measures of signal detectability

Another variant of the MSDR has found application to detection/discrimination experiments. Denoted  $d'_e$ , it gives a measure of the distance between the mean responses to two different signals (or one signal and the ‘noise’ created by spontaneous fibre activity), relative to the spread (standard deviation) of the distributions. Hence  $d'_e$  provides a measure of the detectability or discrimination of a signal. Here  $d'_e$  is computed according to the definition given by Equations (1,2) of Kim et al. (1990):

$$d'_e = \frac{(m_1 - m_2)}{0.5(\sigma_1 + \sigma_2)} \quad (3.19)$$

where  $m_1$  and  $\sigma_1$  are the mean number and standard deviation of spikes counted, and  $m_2$  and  $\sigma_2$  are the mean number and standard deviation of spikes counted due to a second stimulus or the fibre’s spontaneous activity.

Plotted in Figure 3-6 are values of  $d'_e$  for a single, high-CF fibre (with a dead time of 1 ms and a spontaneous rate of 60 spikes/s), calculated via (3.4), (3.7) and (3.19) as a function of discharge rate and integration time. Detectability is seen to be a monotonic function of the discharge rate, with the slope increasing with increasing integration time.

Plotted in Figure 3-7 are values of  $d'_e$  for the output of a summing neuron (with an integration time of 5 ms) which receives input from an arbitrary number of identical AN fibres (each having the same properties as used to plot Figure 3-6), calculated as a function of single-fibre discharge rate and number of input fibres. As was the case for a single fibre, detectability increases monotonically with increasing discharge rate. However, increasing the number input AN fibres up to 50 fibres appears to have a greater effect on the detectability than does increasing the integration time up to 20 ms for a single fibre (*cf.* Figure 3-6).

## 3.4 Application of the model to theories of sound coding

In this section, using the measures of the model’s response properties developed in Section 3.3, summing neurons with different parameters are investigated to determine how they may be used in the processing of AN activity. Two key theories of the neural coding of sound are examined which propose that the human brain uses summation of AN activity in performing certain perceptual tasks. From this it is suggested what degree of AN fibre convergence and temporal integration would be useful in performing these tasks and which CN cell types reported in the literature appear to have the required parameters.

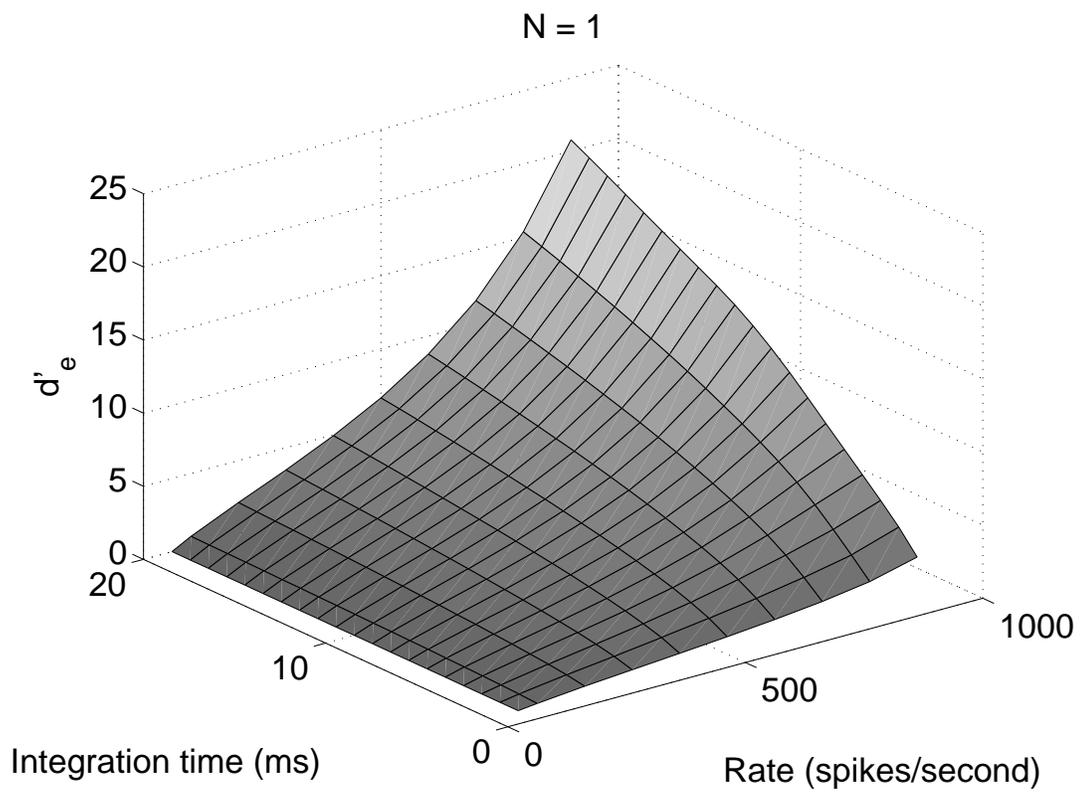


Figure 3-6: Detectability versus discharge rate and integration time.

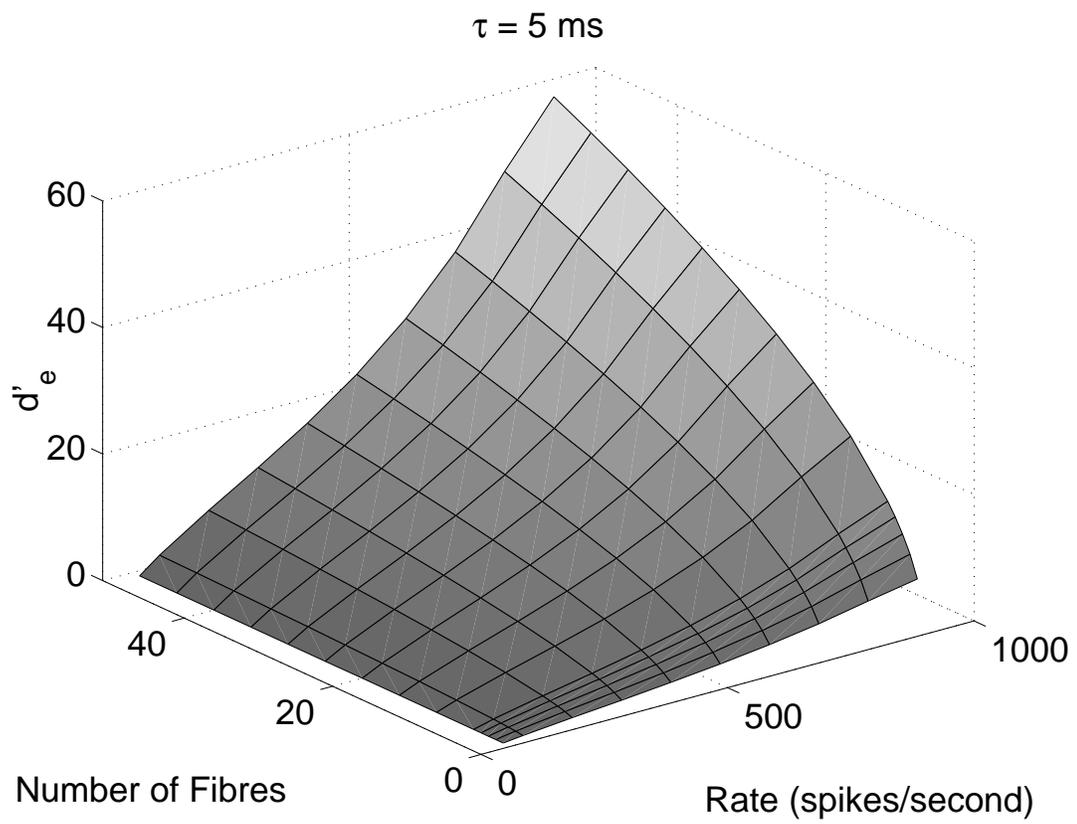


Figure 3-7: Detectability versus single-fibre discharge rate and number of input fibres.

### 3.4.1 Temporal (time-period) frequency coding

Temporal frequency coding is limited by refractory effects, which prevent the coding of periods less than the absolute dead time in the periodicity of single fibre responses. It has been suggested that summation of the response of a number of fibres could overcome this limitation (Wever, 1949). For tone periods shorter than the dead time, the fibres will tend to fire at multiples of the tone period, but due to their stochastic nature each fibre in a population can begin its response at any period of the tone (Clark, 1996; Rose et al., 1967; Wever, 1949). Therefore, if the activity of a population of AN fibres is summed and the sum has a high PSI, and if the summing neuron has a refractory period significantly shorter than that of AN fibres, the tone period will be approximately reconstructed. Furthermore, it has been shown that components of complex acoustical signals are well represented by the average localised synchronised rate (Young and Sachs, 1979) of a population of AN fibres, which could be determined by a summing neuron with a high PSI.

A high PSI will be maintained when the temporal integration time is very short, and when the summation distance is close to zero, minimising the effects of the phase differences in single-fibre responses resulting from the travelling wave and of the reduced synchronisation of fibres distant to the CP. Furthermore, thresholding of the summed response could enhance the synchronisation of the output of a summing neuron. Indeed, there do exist CN cell types which appear to enhance synchronisation through spatiotemporal summation of AN firings (Joris, Carney, Smith and Yin, 1994; Joris, Smith and Yin, 1994). It can be postulated that the AN fibres which converge onto these cells originate from a very narrow region of the basilar membrane, if the total effect of these cells is to enhance synchronisation.

### 3.4.2 Average-rate intensity coding

Figure 3-5 shows that the maximum of the confidence interval for the summing neuron's output is significantly smaller when the PSI is zero than when it is non-zero (even for a relatively small PSI as shown in the lower pane of Figure 3-5), even though the averages of the confidence intervals over time are similar. It is likely that estimation and discrimination tasks based on the average rate of the summed response would need to allow for the maximum range of possible outputs, rather than the instantaneous variance which oscillates over time as a function of the stimulus frequency. Consequently, a non-zero PSI could decrease discrimination and complicate the estimation task.

Long-term integration will reduce the PSI, but will also wipe out any rapid changes in the discharge rate that result from transients in the acoustical signal. An alternative way of estimating the average of an oscillating discharge rate is spatial summation of the activity of a group of fibres, where a phase delay is introduced for each fibre such that the time-varying components cancel out. The basilar membrane mechanics appear to provide such a phase delay. For the analytical model, the PSI can be maintained at 0, independent of the integration time, if the spread of input

fibres corresponds to a  $-2\pi$  phase difference in the synchronised components of the responses (see Figure 3.3(a)). From the computer model (see Figure 3.4(a)), it appears that the loss of synchronisation in fibres distant from the CP causes the spread of fibres needed to obtain a very low PSI to be slightly wider than the first minimum in the analytical model's synchronisation. However, this also causes the synchronisation to remain relatively low for wider spreads of inputs, rather than having a series of pronounced peaks and troughs as predicted by the analytical model. Therefore, if the spread of input fibres is sufficiently wide, the output of the summing neuron will be practically unmodulated, producing more discriminable outputs. Furthermore, it appears from Figures 3-6 and 3-7 that increasing the number of input fibres may create more detectable and discriminable outputs than does increasing the temporal integration time.

In order to gain an appreciation of the spread of inputs required to reduce the synchronisation to close to zero, the spread of inputs is compared with critical bandwidths, a measure of the effective spread of neural excitation which can be obtained by a number of different psychophysical tasks (Moore, 1977). Figure 3-8 shows that the summation distances (o) corresponding to a  $PSI = 0$  are greater than or equal to typical critical bandwidths (\*) reviewed by Moore (1977, p. 103). Thus, the model predicts that the spread of fibres required to produce an output suitable for intensity coding based on the average discharge rate is at least as wide as psychophysically determined critical bandwidths. Consequently, if such summing neurons are involved in intensity coding, they would not cause any spatial filtering of neural excitation narrower than critical bandwidths, and critical bandwidth measurements based on intensity cues should not be affected by the summing neurons presence.

At the threshold of hearing, the area of neural excitation will be significantly smaller than the spread of fibres required to maintain low synchronisation at higher intensities (see the top section of Figure 3-2), meaning that the summing neuron's output *will* be modulated. However, at low intensity levels, detection of the acoustical signal is of importance, rather than estimation of the signal intensity or the form of the percept. In this case, it could be desirable for the summed response to be modulated, because this increases the chances of the summed response reaching the neuron's threshold at some stage. Such behaviour, known as *stochastic resonance*, has been observed in a range of natural systems and has been determined to be useful in the detection of weak signals (Moss and Wiesenfeld, 1995). Of particular interest is that summing neurons can make optimal use of stochastic resonance more easily than can single neurons (Collins et al., 1995).

From these two results it can be hypothesised that one type of spatiotemporal summing neuron (wide spread of input AN fibres, short temporal-integration window) could be used both for detection of sounds at low intensity levels and for estimation of the sound intensity at higher intensity levels.

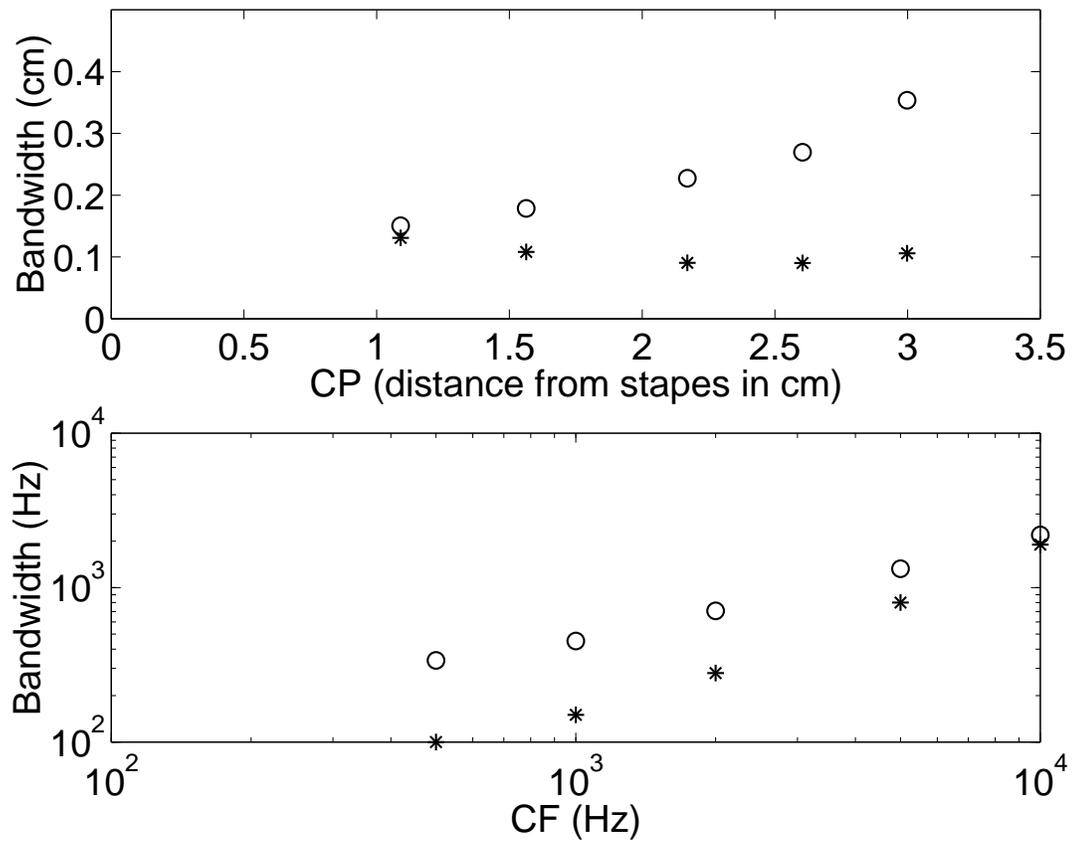


Figure 3-8: Zero-PSI widths (o) and typical critical bandwidths (\*) from a review by Moore (1977, p. 103). Top: Widths versus CP; Bottom: Widths (expressed in terms of the corresponding difference in CF) versus CF.

### 3.5 Discussion

The analytical models developed in this chapter permit quantitative analysis of the effects of different degrees of AN fibre convergence onto summing neurons in the CN and of possible uses of such neurons in the processing of AN activity. The modelling results suggest that CN cells which show enhanced synchronisation may need to receive input from AN fibres originating from sites on the basilar membrane which are very close together. This is primarily because of the effect of the phase differences in single-fibre responses resulting from the wave. A second factor is the reduced synchronisation of fibres distant from the CP. The results also suggest that the degradation in synchronisation found in other CN cells could actually be a result of a wide spread of input fibres, not necessarily a long temporal integration time-constant or some other physiological mechanism. Furthermore, a near-to-total degradation of synchronisation resulting from a wide spread of inputs may actually be a useful way of determining the average rate of AN fibres in a population without requiring long-term temporal integration, which would wipe out information coded by rapid changes in the instantaneous rate.

These results prompt further anatomical and physiological investigation of which AN fibres converge onto different types of summing neurons in the CN. In particular, the results prompt a search for the existence of summing neurons which reduce synchronisation by taking inputs from a wide spread of AN fibres. The impetus for such an investigation might not have resulted from direct analysis of physiological data, because (i) reduced synchronisation is possibly less conspicuous than enhanced synchronisation, and (ii) a function might not have been so readily postulated for such behaviour.

The results also indicate the ability of models such as the one developed in this chapter to provide new insights into the physiology of the auditory system. One possible extension to the model would be to investigate the effects of different weightings for the inputs to the summing-neuron model, such as a graded spreading function (e.g., closer sources have greater weights), a statistical spread of influence, or adapted patterns of sensitivity as seen in mature neural systems. Another useful extension would be to derive expressions for the output of the summing neuron with a thresholding mechanism added, using a method such as that developed by Burkitt and Clark (1997). This would allow quantitative analysis of the interaction of the reduction of synchronisation produced by the spread of AN inputs, as shown in this chapter, with the enhancement of synchronisation produced by thresholding (Rothman and Young, 1996; Rothman et al., 1993; Burkitt and Clark, 1997).

## Chapter 4

# The effect of neural thresholds on frequency estimation

### 4.1 Introduction

Auditory nerve fibres differ in their spontaneous firing rates, and in their response thresholds. These response differences suggest that auditory sound coding could be based on more than just the characteristic frequency of the neurons. In fact, physiological experiments demonstrate that when stimulated by a complex sound, fibres with low spontaneous rates predominantly respond to the envelope, and those with high spontaneous rates to the fine temporal structure of the sound (Langner, 1992; Horst et al., 1986).

The neural response thresholds are highly correlated to the spontaneous rates (Winter et al., 1990; Liberman, 1978), but reasonably independent of characteristic frequency. Thus the brainstem receives information from fibres which may approximately be parameterised in a two-dimensional response space—where one parameter represents characteristic frequency, and the other represents threshold. Much work has been done to understand how the responses of fibres with different characteristic frequencies are synthesised for the task of frequency estimation (Kim and Parham, 1991; Miller and Sachs, 1984; Sruлович and Goldstein, 1983), but very little analysis has been applied to understanding the role that fibres with different thresholds play in the same task.

In this chapter the effects of filtering by a multi-threshold system on the estimation of complex acoustic stimuli are investigated. This is achieved by generating a model of neural response to a complex sound, and investigating via Cramer–Rao bounds (Kay, 1993), and via the distribution of interspike intervals, the accuracy to which information about the frequencies within the sound may be estimated, based either on the output of high-threshold and/or low-threshold fibres.

## 4.2 Signal and network models

### 4.2.1 Signal model

Consider a common estimation task performed by the auditory system: the analysis of the frequency components of a speech signal. Such signals are composed of complex sounds which exhibit a number of resonances (formants), all modulated by a voicing pitch. Perceptual experiments show that if the actual fundamental of the voicing pitch is missing from the spectrum, then the estimated voice pitch corresponds to the smallest difference between the harmonics present. This is a well noted auditory phenomenon known as the ‘missing fundamental’ (Langner, 1992; Javel and Mott, 1988).

Here the estimation of the voice pitch from the temporal characteristics of neural response is explored for a signal where two harmonics of the voice pitch are present, but not the fundamental.

The input to the neural network is taken to be the sound pressure wave passed through a linear filter, the cochlea. The filter characteristics of the cochlea to a 700 Hz tone is shown in Figure 4-1.

Thus, the filtered signal  $s(t)$  is expressed as:

$$s(t) = 1 + \sum_{i=1}^2 A_i \sin(2\pi f_i t + \phi_i) \quad (4.1)$$

where  $f_1$  and  $f_2$  are the harmonic components of the voice pitch present in the filtered signal.

The task is to estimate the voice pitch,  $f_2 - f_1$ , and its harmonics,  $f_2$  and  $f_1$ . This could be done either by estimating  $f_1$  and  $f_2$  simultaneously and calculating the difference, or by introducing the  $f_2 - f_1$  component to the signal via a nonlinearity and estimating the voice pitch directly from the modified signal.

### 4.2.2 Neural network model

The response of neurons of the auditory nerve may be modelled by an inhomogeneous Poisson process (Siebert, 1970), where the instantaneous expected rate is described by means of a compressively-nonlinear (sigmoidal) function, which is brought about by a number of nonlinearities involved in AP thresholding and generation. From the form of input-output curves derived from physiological data (Sachs et al., 1989),  $\tanh(\cdot)$  is taken to be a suitable sigmoidal function. Although spontaneous rate is routinely used to classify fibre responses, a threshold shift can better explain the differing responses (Horst et al., 1990).

Thus the Poisson rate of the  $n^{th}$  neuron,  $r_n(t)$ , may be described by:

$$r_n(t) = r_0 + \tanh(\alpha_n [s(t) - \beta_n]) \quad (4.2)$$

where  $s(t)$  is the cochlear filtered signal defined in the previous subsection.

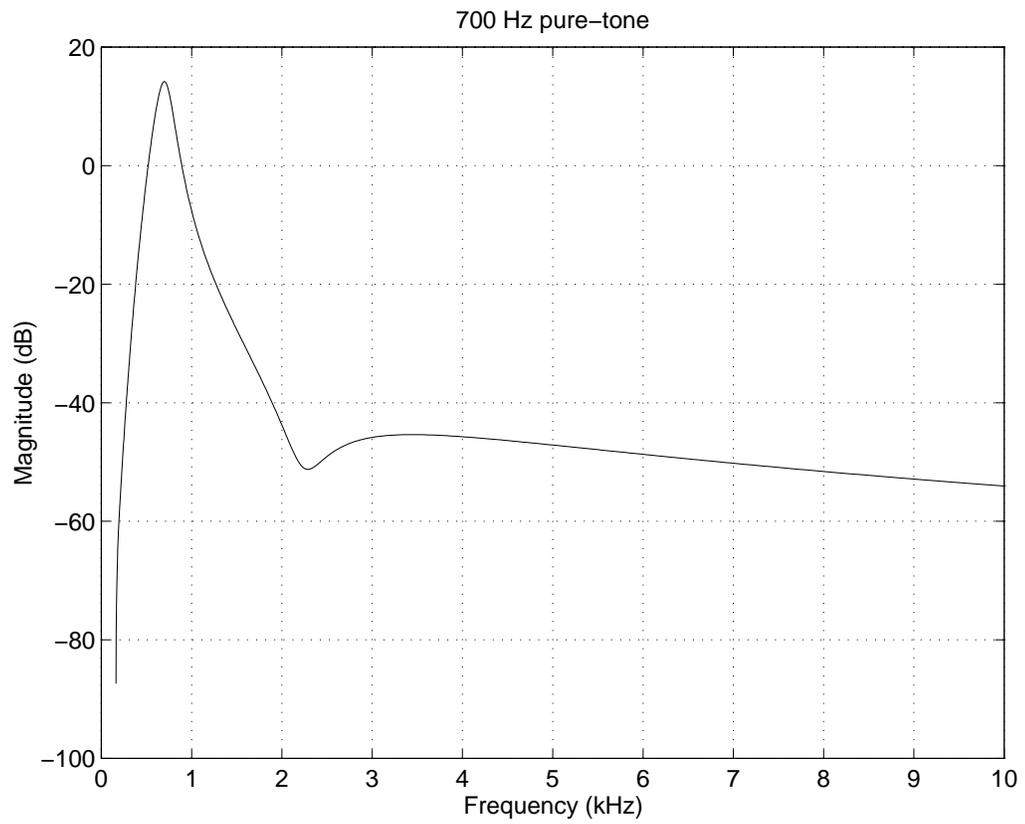


Figure 4-1: Filter characteristics of the cochlea to a 700 Hz tone.

### 4.2.3 Filtering properties of the neural model

Changing the steepness and position of the sigmoidal  $\tanh(\cdot)$  allows the simulation of a range of neural responses with various onset and saturation thresholds, and these nonlinear responses will attenuate or magnify various components of the sound spectrum. Fourier analysis can be used to find the threshold value that minimises a cost function which measures the relative magnitude of a specific frequency component at the output of the sigmoid. Such a cost function can include a penalty function which prevents the absolute magnitudes of the major components from being overly attenuated.

For the estimation task described in Section 4.2.1, the components of interest are  $f_1$  and  $f_2$ , and the missing fundamental of the voice pitch,  $f_2 - f_1$ . Consequently, the analysis described above is performed to determine thresholds which accentuate each of these components.

The relative magnitudes of the two voice-pitch harmonics present in  $s(t)$  will depend on their magnitudes in the sound pressure wave and on the filter characteristics of the cochlea at the place of the fibre's input. It is therefore possible to have a range of modulation depths in the signal. Here two signals are investigated with magnitudes chosen arbitrarily to produce a slightly modulated  $s(t)$  (Example 1) and a highly modulated  $s(t)$  (Example 2).

#### Example 1: Slightly modulated $s(t)$

Consider the signal and neural response:

$$\begin{aligned} s(t) &= 1 + \frac{1}{6} \sin(2\pi 600t) + \frac{5}{6} \sin(2\pi 700t) \\ r(t) &= 1 + \tanh(10(s(t) - T)) \end{aligned}$$

where  $T$  is a threshold shift.

#### Example 2: Highly modulated $s(t)$

Consider a neural response the same as for Example 1, but with the signal:

$$s(t) = 1 + \frac{1}{2} \sin(2\pi 600t) + \frac{1}{2} \sin(2\pi 700t)$$

For both examples, Fourier analysis of the output of each sigmoid was used to maximise the relative size of its components at the frequencies 600, 700 and 100 Hz from among the parameterised sigmoid function:

$$r(t) = 1 + \tanh(10(s(t) - T))$$

The optimal threshold values,  $T$ , are shown in Table 4.1.

Table 4.1: Optimal thresholds for Examples 1 and 2

Frequency (Hz)	600	700	100
Ex. 1: Optimal T	1.73	1.00	1.75
Ex. 2: Optimal T	0.97	1.00	1.68

The sigmoids are shown in Figure 4-2, and the input and output of the sigmoids and their Fourier transforms are shown in Figures 4-3 and 4-4. The implications of the filtering properties of the sigmoidal nonlinearity will be investigated in the next section.

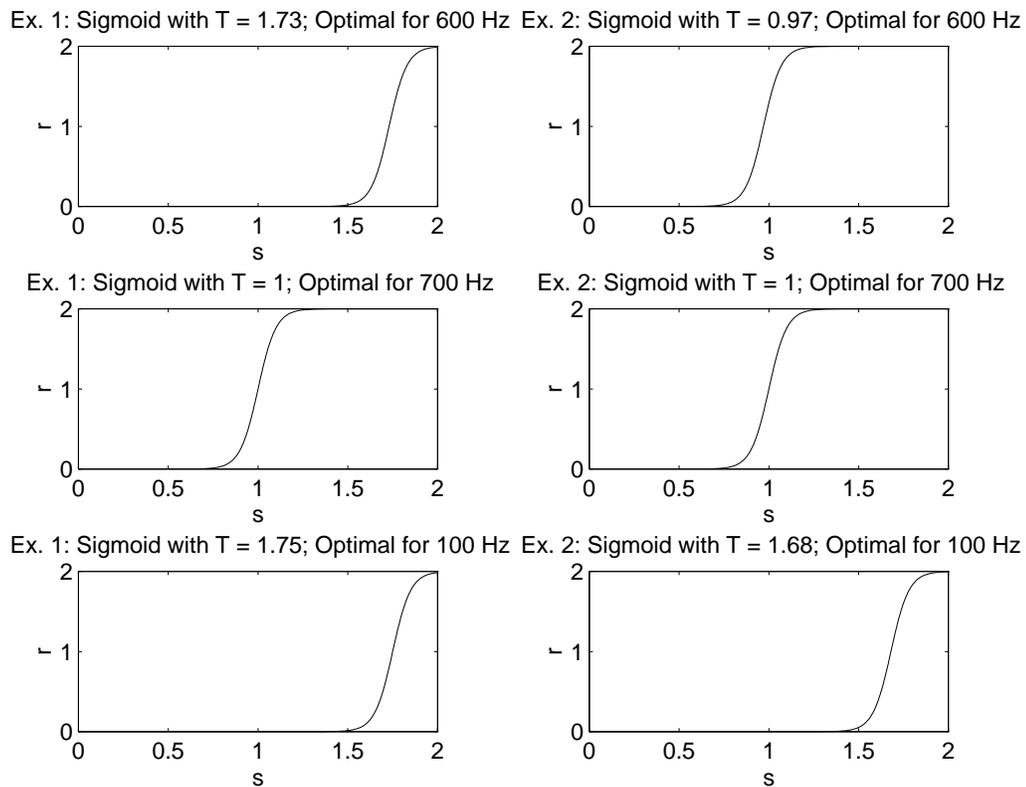


Figure 4-2: Left: Ex. 1—Slightly modulated  $s(t)$ . Sigmoids with optimal thresholds for 600, 700 and 100 Hz (top to bottom). Right: Ex. 2—Highly modulated  $s(t)$ . Sigmoids with optimal thresholds for 600, 700 and 100 Hz (top to bottom).

#### 4.2.4 Cramer–Rao bounds for neural estimation of frequency

The auditory system takes the responses of some 30 000 auditory nerve neurons, and can produce estimates of the amplitudes,  $A_i$ , and frequencies  $f_i$  of the sound  $s(t)$ . Exactly how this is achieved is largely unknown, however statistical methods can yield information about the ability of any

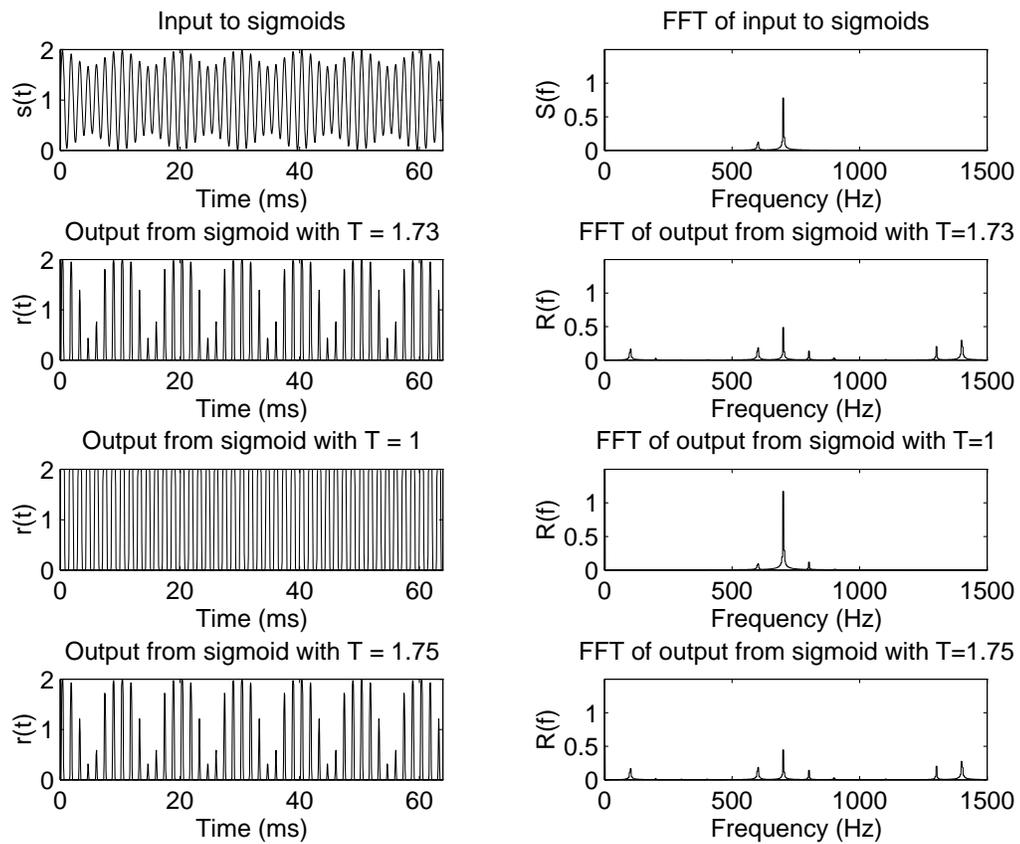


Figure 4-3: Ex. 1: Slightly modulated  $s(t)$ . Output values of the sigmoids with optimal thresholds for 600, 700 and 100 Hz.

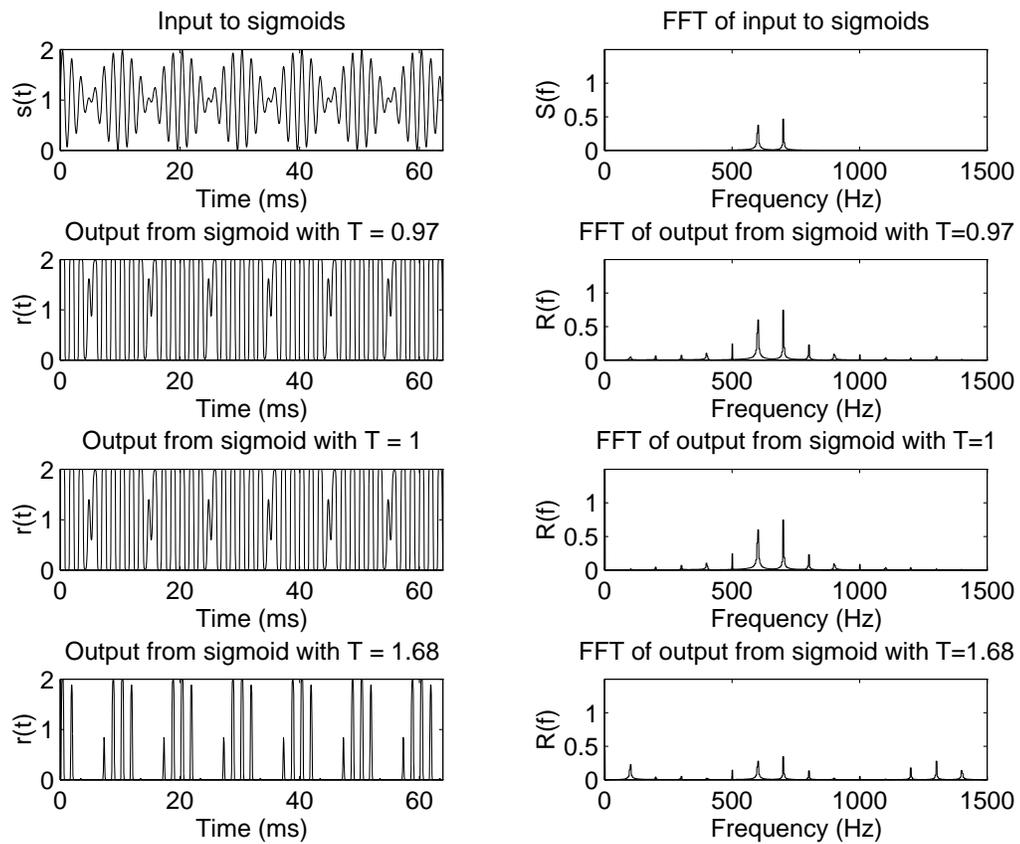


Figure 4-4: Ex. 2: Highly modulated  $s(t)$ . Output values of the sigmoids with optimal thresholds for 600, 700 and 100 Hz.

proposed neural structures to estimate properties of the sound. In turn, these abilities help shed light on likely mechanisms for the information processing capabilities of the auditory system.

One method of analysing the ability of proposed mechanisms to code parameters is via the application of the Cramer–Rao Bound (Kay, 1993). This permits a lower-bound to be given for the variance of any unbiased estimator for the parameter in question. Of course, such an optimal estimator may not exist, or even be compatible with the structures of the auditory system. Such an analysis is still useful, however, because it can rule out mechanisms which do not convey the required information.

The following Lemma is based on calculations performed in Siebert (1970), for the estimation of a pure tone.

**Lemma 4.2.1** Consider an observation for duration  $\mathcal{T}$  of an inhomogeneous Poisson process with rate  $r(t, f, A)$ . Then the Cramer–Rao inequality can be expressed as:

$$\frac{1}{\hat{\sigma}_f^2} \leq \int_0^{\mathcal{T}} \frac{1}{r(t, f, A)} \left[ \frac{\partial r(t, f, A)}{\partial f} \right]^2 dt$$

This result can be extended to define the Fisher Information Matrix  $\mathbf{I}(\boldsymbol{\theta})$ , for the estimation of the unknown vector parameter  $\boldsymbol{\theta} = [A_1, f_1, \phi_1, A_2, f_2, \phi_2]'$ , where  $A_i, f_i, \phi_i$  are the parameters of  $s(t)$  as described in (4.1).

**Lemma 4.2.2** Consider observations of a number of inhomogeneous Poisson processes, with rates  $r_n(t, \boldsymbol{\theta})$ . In this case, the Cramer–Rao inequality can be expressed as:

$$\begin{aligned} [\mathbf{I}_n(\boldsymbol{\theta})]_{ij} &= \int_0^{\mathcal{T}} \frac{1}{r_n(t, \boldsymbol{\theta})} \frac{\partial r_n(t, \boldsymbol{\theta})}{\partial \theta_i} \frac{\partial r_n(t, \boldsymbol{\theta})}{\partial \theta_j} dt \\ \sigma_n^2(\theta_i) &\geq [\mathbf{I}_n^{-1}]_{ii} \end{aligned} \tag{4.3}$$

**Remark 4.2.1** A standard result of Cramer–Rao theory, shows that the information matrix of the combined results of independent experiments equals the sum of the information matrices of each individual experiment. Thus, under the assumption of conditional independence of auditory nerve responses, a calculation of the Fisher Information Matrix of the output of two or more neurons can be achieved by summing the individual matrices. This facilitates easy comparison of the output of various groups of fibres, and the ability to take the output of one fibre, and select the fibre which minimises the estimator variances based on the combined information of both fibres.

Thus, the evaluation of Lemma 4.2.2 where the response rates are taken from the signal and neural models of Sections 4.2.1 and 4.2.2 enables calculation of bounds on estimator performance based on the outputs of a number of neurons.

In the case of a sigmoidal response function (4.2), and sinusoidal signal model (4.1), the integral of (4.3) does not appear to be analytically tractable, and consequently it is not solvable for generalised conditions. However, it is numerically solvable for any given parameters, and in a later section the estimator variance is investigated for some representative situations.

### 4.2.5 Interspike interval analysis

Although it is still unknown exactly what mechanisms exist by which the auditory system codes frequency (Clark, 1996), it has been hypothesised that one method may be via the detection of dominant intervals between neural responses—effectively picking the period of the response waveform. This could be achieved via a series of delay lines and coincidence detectors (Clark et al., 1996; Licklider, 1954). How would the thresholding neurons effect this kind of system?

Although the Cramer–Rao bounds of the previous sections can limit the variance of estimators based on neural responses, they can not indicate the degree to which the auditory system’s variances follow the optimal bounds, and consequently are not necessarily an accurate measure of how useful the output of a selected neuron is to the auditory system. To investigate this question, the distribution of interspike intervals is used.

**Lemma 4.2.3** Consider an Inhomogeneous Poisson Process with rate  $s(t)$ , over the time interval  $[0, \mathcal{T}]$ . Then the distribution of spikes occurring with a gap of  $\tau$  is:

$$D(\tau) = \frac{\int_0^{\mathcal{T}-\tau} s(t)s(t+\tau)dt}{\left[\int_0^{\mathcal{T}} s(t)dt\right]^2 / 2}$$

**Remark 4.2.2** This distribution measures the relative frequency of spikes occurring with a time difference of  $\tau$ , regardless of the existence of spikes within the interval. It is consistent with the type of estimator proposed earlier in this section. An alternative expression for the distribution of intervals for *successive* spikes only may also be derived (see Chapter 2).

The results of Lemma 4.2.3 are used in a later section to calculate the effect of the sigmoidal nonlinearity on the interspike-interval distribution. Similar to the Cramer–Rao bounds, the integral appears analytically intractable, but can easily be calculated numerically for specific examples.

## 4.3 Results

In this section, the effect of various neural thresholds on the frequency estimation task of Section 4.2.1 is determined for the case of an efficient estimator, and also for an interspike-interval

based estimator.

### 4.3.1 Estimation by an efficient estimator

Analytical descriptions of the inner terms of the integral given in (4.3) are derived for a signal of the form expressed in (4.1). The integrals, however, appear not to be analytically tractable, and therefore numerical integration was implemented using an adaptive recursive Newton Cotes 8 panel rule. Cramer–Rao Bounds were calculated for Examples 1 and 2 of Section 4.2.3, with all three possible combinations of low-threshold ( $T = 1.00$ ) and high-threshold ( $T = 1.80$ ) sigmoid pairs (L+L, H+H and L+H). The bounds were evaluated over 20 ms (Table 4.2) and 100 ms (Table 4.3).

Table 4.2: Cramer–Rao Bounds for Examples 1 and 2: 20 ms

Ex.	T	600 Hz	700 Hz	100 Hz	$\Sigma$
1	L+L	7.24e-3	1.73e-4	7.80e-3	1.52e-2
	H+H	4.32e-3	3.77e-4	5.05e-3	9.75e-3
	L+H	5.38e-3	2.31e-4	5.97e-3	1.16e-2
2	L+L	4.01e-4	4.03e-4	9.00e-4	1.70e-3
	H+H	1.23e-3	1.16e-3	2.65e-3	5.04e-3
	L+H	5.89e-4	5.82e-4	1.34e-3	2.51e-3

Table 4.3: Cramer–Rao Bounds for Examples 1 and 2: 100 ms

Ex.	T	600 Hz	700 Hz	100 Hz	$\Sigma$
1	L+L	5.59e-5	1.16e-6	5.72e-5	1.14e-4
	H+H	2.85e-5	3.04e-6	3.16e-5	6.31e-5
	L+H	3.77e-5	1.67e-6	3.95e-5	7.89e-5
2	L+L	2.85e-6	2.88e-6	5.46e-6	1.12e-5
	H+H	8.32e-6	8.27e-6	1.69e-5	3.35e-5
	L+H	4.23e-6	4.27e-6	8.24e-6	1.67e-5

In some cases the low-threshold sigmoids produced the smallest bound, in others the high-threshold. In none of the examples studied did the combination of thresholds (L+H) produce the smallest error, however when averaged over the signal models studied the L+H combination had a lower mean error (20 ms: 7.05e-3; 100 ms: 4.78e-5) compared to the L+L (20 ms: 8.45e-3; 100 ms: 6.25e-5) and H+H (20 ms: 7.40e-3; 100 ms: 4.83e-5) combinations. Thus for the

particular situations investigated, an ‘efficient’ estimator will operate best when the thresholds are identical, however different signals will require different optimal thresholds, validating the need for auditory nerve fibres with a range of thresholds.

### 4.3.2 Estimation from the interspike-interval distribution

One possible mechanism for frequency estimation would be to utilise the interspike-interval distribution as examined in Section 4.2.5 to measure the dominant interspike interval observed between two fibres. Figure 4-5 shows the relative frequency of occurrence of interspike intervals for Examples 1 and 2, with a low-threshold and a high-threshold sigmoid.

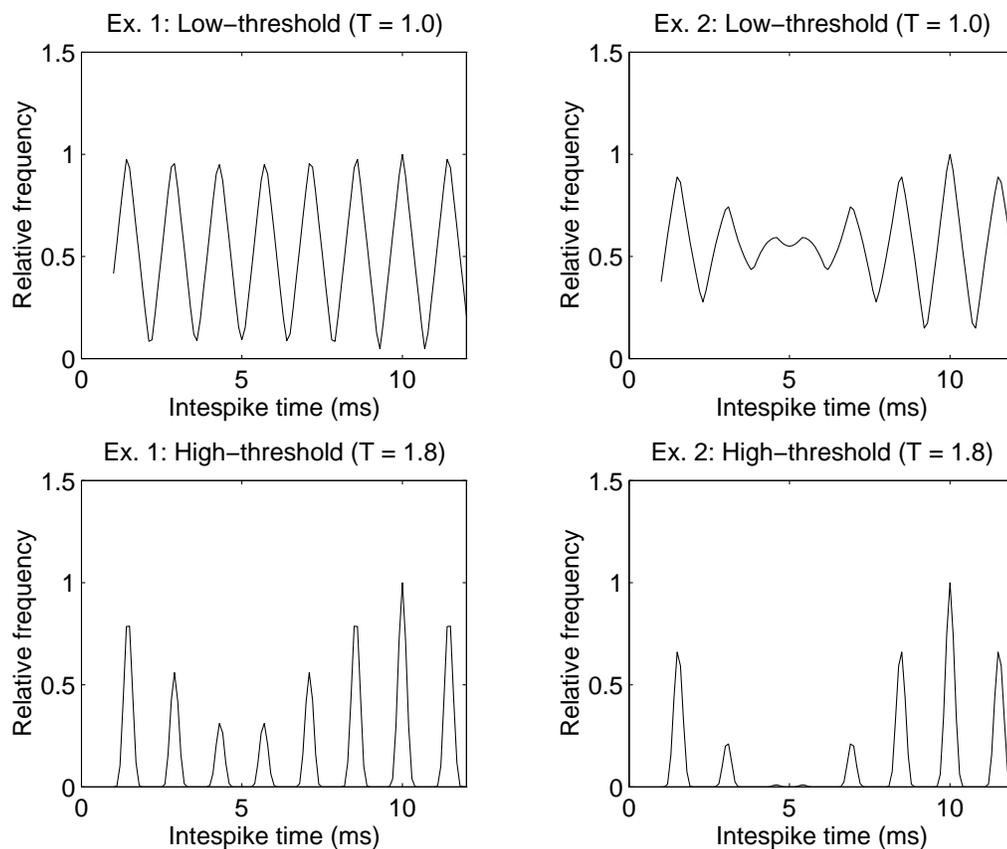


Figure 4-5: Left: Ex. 1—Slightly modulated  $s(t)$ . Interspike-interval distribution for low-threshold ( $T = 1.0$ ) and high-threshold ( $T = 1.8$ ) sigmoids (top and bottom). Right: Ex. 2—Highly modulated  $s(t)$ . Interspike-interval distribution for low-threshold ( $T = 1.0$ ) and high-threshold ( $T = 1.8$ ) sigmoids (top and bottom).

For both Examples 1 and 2, the interspike interval corresponding to the weighted average of the periods resulting from the 600 and 700 Hz components of the signal (1.47 ms for Ex. 1; 1.55 ms for Ex. 2) is emphasised by the low-threshold sigmoids, and to the period of the 100 Hz ‘missing fundamental’ (= 10 ms) by the high-threshold sigmoid. This suggests that for a system

measuring interspike intervals, a combination of low-threshold and high-threshold fibres is useful to estimate all three frequency components when using this estimation technique, particularly if further filtering is used to extract only one frequency per fibre (e.g., Srulovicz and Goldstein, 1983).

## 4.4 Discussion

Auditory nerve fibres may be parameterised in terms of their response properties including the frequency of sound to which they best respond, and their response thresholds. For the task of estimation the different frequency components of complex stimuli, the importance of combining the output of auditory nerve fibres with differing thresholds was measured. The resulting Cramer–Rao bounds permit the calculation of frequency estimation variance for any given neural parameters and show, for the calculated examples, that an ‘efficient’ observer of the output of two fibres may benefit from a mixing of differently-thresholded neurons. It is also demonstrated how combining the output of differently-thresholded fibres may improve frequency estimation based on interspike intervals. Due to the numerical nature of the calculations, these results are not calculated parametrically, but are demonstrated for a number of specific examples. An open question is the extension of these results to a general case—thereby specifying conditions under which the combined threshold responses are more useful than single threshold responses, and conditions under which they are not.

## Chapter 5

# A stochastic model of the electrically stimulated auditory nerve: Single-pulse response

### 5.1 Introduction

In the previous chapters it has been investigated how spatiotemporal combination of auditory nerve activity may be used in the processing and decoding of acoustical stimulus features. In the case of *electrical* stimulation, spatiotemporal combination of auditory nerve activity has also been postulated to explain why single-fibre physiological data have been unable to explain some important psychophysical phenomena observed in cochlear implant users (e.g., Pfingst et al., 1991). This hypothesised effect of spatiotemporal combination of auditory nerve activity may require that stochastic components of the single-fibre response to electrical stimulation be considered (White, 1984a).

Historically, neural response to electrical stimulation has been modelled via deterministic descriptions (exceptions among cochlear implant groups include O’Leary et al. (1998), Irlicht and Clark (1996), Dynes (1996), Rubinstein (1995), White et al. (1987), Hochmair-Desoyer et al. (1984) and White (1978); and in other fields of neuroscience Clay and DeFelice (1983) and Lecar and Nossal (1971)). However, a small amount of stochastic activity was present in the squid giant axon (Hodgkin and Huxley, 1952) and frog node (Dodge, 1961; Frankenhaeuser, 1957) data on which many deterministic models are based (Verveen and Derksen, 1965). Furthermore, significant variance has been measured in the response of nerve fibres to single current pulses (Dynes, 1996; Hoopen and Verveen, 1963; Verveen, 1961; Verveen, 1960) and pulse trains (O’Leary et al., 1998), which cannot be explained by deterministic models.

However, the amount of variance *is* dependent on the stimulus intensity. It has been generally assumed that cochlear implants operate at stimulus intensities which produce extremely high discharge probabilities (very close or equal to 1) in AN fibres, resulting in very little variance. In contrast, preliminary evidence is emerging that cochlear implants may be operating at intensities which produce low to moderate discharge probabilities, i.e., significant variance (see Chapter 6). If this is correct, then there are important consequences for cochlear implants. For example, (i) both temporal and spatial aspects of speech coding would be impacted (see Chapters 6 and 7), and (ii) the survival of AN fibres subject to high stimulation rates is likely to be strongly affected (Xu et al., 1997; Tykocinski et al., 1995c).

Verveen and Derksen (1968) showed that the variance in response to single current pulses can be attributed to random fluctuations in the voltage across the membrane of myelinated nerve fibres at the nodes of Ranvier. They recorded membrane-potentials *in vitro* from sciatic nerve of the frog, *Rana temporaria*. Figure 5-1 shows a number of their membrane-potential traces at  $-90$  mV 5.1(a) and at a range of depolarised and hyperpolarised potentials 5.1(b) (resting potential is  $\sim -70$  mV (Derksen and Verveen, 1966)).

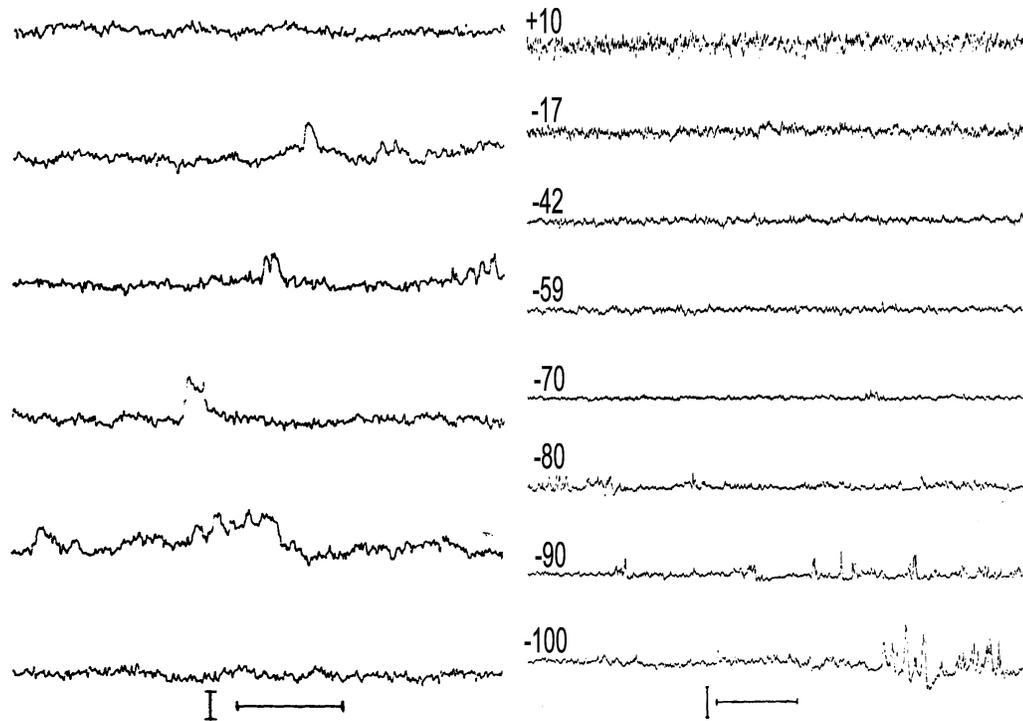
Verveen and Derksen (1968) characterised these fluctuations as having a Gaussian amplitude distribution and a  $1/f$  frequency spectrum<sup>1</sup>. In the upper traces of Figure 5.1(b) it appears that the amplitude of the fluctuations increases as the membrane is depolarised. Furthermore, some non-stationary behaviour is seen in the middle traces of Figure 5.1(a) and in the lower traces of Figure 5.1(b). These non-stationary fluctuations appear to increase both in frequency of occurrence and in amplitude with increasing hyperpolarisation.

Membrane potential fluctuations in squid axons are considered to be of negligible magnitude (Verveen and Derksen, 1965). However, Verveen (1962) reported that the magnitude of the membrane potential fluctuations appears to be negatively correlated with the diameter of the nerve fibre. Mammalian auditory nerve fibres have a mean diameter of  $\sim 2-4$   $\mu\text{m}$  (Gleich and Wilson, 1993; Spendlin and Schrott, 1989) and even smaller diameters ( $0.1-0.7$   $\mu\text{m}$ ) at the nodes of Ranvier in the myelinated portion of the peripheral dendrite, proximal to the habenula (Lieberman and Oliver, 1984). This is two to three orders of magnitude smaller than the mean diameter of squid giant axons,  $\sim 500$   $\mu\text{m}$  (Hodgkin, 1961), which indicates why membrane potential fluctuations may be of significant magnitude in mammalian auditory nerve fibres.

Further physiological (Sigworth, 1980; Derksen and Verveen, 1966) and modelling (Rubinstein, 1995) studies suggest that these voltage fluctuations are likely to be caused by random transitions in the state of single ionic channels at the nodes of Ranvier. This is consistent with the magnitude of the fluctuations being negatively correlated with the diameter of the axon, since the larger the

---

<sup>1</sup>A  $1/f$  frequency spectrum indicates that, while the distribution of amplitudes is Gaussian, auto-correlation exists between these amplitudes, i.e., the amplitude tends to change slowly over time without affecting the relative frequency of occurrence of amplitudes over an extended period of time.



(a) Membrane-potential traces at  $-90$  mV. Horizontal scale: 10 ms; vertical scale: 1 mV. Reprinted with permission from Figure 20 of Verveen and Derksen (1968) © 1968 IEEE)

(b) Membrane-potential traces at different levels of membrane potential (given in mV above each trace). Horizontal scale: 1 s; vertical scale: 5 mV. Reprinted with permission from Figure 19 of Verveen and Derksen (1968) © 1968 IEEE)

Figure 5-1: Membrane-potential traces from Verveen and Derksen (1968) showing fluctuations in nerve fibre transmembrane potentials at the nodes of Ranvier.

diameter of the axon, the more single ionic channels there will be at each node. The more channels, the less significant the fluctuations of any one channel will be when compared with the total ionic conductance. Indeed, introduction of the correct single-channel conductance statistics to models such as the Hodgkin–Huxley equations has produced the correct prediction of membrane potential fluctuations and the resulting response to single pulses and pulse trains (Rubinstein, 1995; Lecar and Nossal, 1971).

However, detailed stochastic descriptions of AN response to electrical stimulation based on Hodgkin–Huxley type models (Dynes, 1996; Rubinstein, 1995; Hochmair-Desoyer et al., 1984) are too computationally intensive to permit simulation of a large number of neurons over hundreds of milliseconds, which may be required for the simulation to be psychophysically relevant. An alternative statistical approach is to apply a point process description of neural response (Mark and Miller, 1992; Miller and Mark, 1992; Bi, 1989; Jones and Tubis, 1985; Johnson and Swami, 1983; Gaumond et al., 1982; Miller, 1971; Perkel et al., 1967) to electrically elicited responses (O’Leary et al., 1998; Irlicht and Clark, 1996). This description is more computationally efficient, but since it is entirely phenomenological it is not easily related to stimulus and neural parameters. Furthermore, (i) it is not clear how this model specialises to an equivalent deterministic model, and (ii) the point process description is not easily generalised to a population of fibres.

In this chapter, an alternative model is developed of AN response to electrical stimulation, following the conceptual approach used in White et al. (1987), White (1978), and Verveen and Derksen (1968). Based on the Hill (1936) threshold model, the model includes a number of significant components of action potential generation, including membrane noise as recorded by Verveen and colleagues. Threshold models are much simpler conceptually and are more computationally efficient than Hodgkin–Huxley type models. They have also been shown to provide a good approximation to the more complex models (Kistler et al., 1997). Additionally, the model can be fitted easily to the statistics of AN parameters collected from physiological studies. It is therefore suitable for the calculation of large-scale population responses. Such responses are required for the investigation of sound coding in ensembles of nerve fibres, for the explanation or prediction of psychophysical results, or for the development of speech processing strategies for cochlear implants.

The Methods section describes the analysis of the physiological data and the methods used to derive the neural model. In Section 5.3 the model of AN response to single biphasic electrical pulses is developed. Response properties of the model are investigated. In particular, comparisons are made between the behaviour of the model with and without the stochastic component. In Section 5.4, the model’s suitability for describing large-scale population responses is demonstrated by developing a model of total AN response. It is shown how the parameters of an arbitrary number of AN neurons can be fitted to the statistics of AN data collected from physiological studies. Using a simple model of current spread it is shown how the output of the total AN model

varies with stimulus and model parameters. Finally, in the Discussion the implications of the stochastic behaviour of AN response for cochlear implant research are explored.

## 5.2 Methods

### 5.2.1 Physiological data

The first set of physiological data presented in this chapter are single-fibre cat AN data from Javel et al. (1987). A previously unpublished analysis of this data set is presented here, in which the probability of discharge in response to single biphasic current pulses is estimated.

The probability of discharge for a single biphasic pulse is estimated from the fraction of trials in which an action potential is generated in response to the first pulse of a pulse train. Because the first pulse of each pulse-train occurs at least 50 ms after the end of the preceding pulse-train, the response to the first pulse should closely approximate the response to a single pulse in isolation<sup>2</sup>. Data of the same stimulus intensity and pulse width (phase duration) are pooled across pulse trains of different rates for each neuron.

The method used in suppressing the stimulus artefact meant that little data was collected for stimuli with pulse widths above 400  $\mu\text{s}$ /phase and none above 600  $\mu\text{s}$ /phase. Pulse width has a significant effect in a range of psychophysical measures and it would therefore be useful for the model to be accurate over a greater range of pulse widths. Accordingly, supplementary physiological data, also from single AN fibres of cats, were collected by Dynes (Bruce, Irlicht, White, O’Leary, Dynes, Javel and Clark, 1997a).

For this second set of physiological data, animal preparation and techniques to record from single AN fibres are essentially as described in Kiang et al. (1965). In brief, adult cats were injected peritoneally with daillyl barbituric acid in urethane solution (dosage 75 mg per kg weight). A cannula was inserted into the trachea, the cartilaginous external auditory meatus was cut near the tympanic membrane, and the bulla opened to expose the round window.

The posterior fossa of the cranial cavity was opened dorsally and the cerebellum retracted medially to expose the auditory nerve. The animal was placed in a sound-proof, vibration-isolated, electrically-shielded chamber. An acoustical cavity containing a calibrated sound transducer (Radio Shack model 40-1377) was sealed in the meatus. Neural activity was recorded through micropipettes (FHC #30-31-0) pulled to a resistance of 15-20 Mohms and filled with 2M KCl and placed in the nerve under visual control.

Electrical stimulation was provided through a single Teflon-coated platinum wire (0.005 in wire diameter). The apical tip of platinum wire electrodes passed about 4–5 mm through the round window. In all cases, the indifferent electrode was an 18-gauge needle placed in the ipsilateral

---

<sup>2</sup>See Chapter 7 for an investigation of inter-pulse interactions.

forelimb.

The electrical search stimulus was a train of 100  $\mu$ sec cathodic pulses presented at a rate of 10/sec. The search stimulus amplitude was varied, but was typically 15 dB below 1 mA. Single units were detected by monitoring the neural potentials on an oscilloscope. The experimental stimulus consisted of symmetric anodic/cathodic biphasic stimuli of differing durations, presented at a rate such that there was at least an 80 msec interval between stimuli. The determination of the probability of discharge for each stimulus level was based on 100 stimulus presentations.

### 5.2.2 Model of AN response

The various single-fibre models are defined in the form of electrical circuits. Monte Carlo simulations of these circuits can be carried out on a computer and the response of the model recorded directly. This method requires the least number of simplifying assumptions and approximations, producing the most accurate description of the model's response properties, but is computationally laborious. This is because the stochastic version of the model requires many simulation iterations to accurately describe the distribution of responses.

In order to overcome this problem, appropriate simplifying assumptions and approximations are made, enabling the derivation of analytical approximations of the electrical circuit descriptions. These assumptions and approximations are tested via Monte Carlo simulations and comparisons with physiological data and are found to be reasonable and accurate. Predictions of single-fibre response can be computed simply and efficiently using these analytical approximations and multiple iterations are not required.

## 5.3 Single-pulse response

Investigating the response of a single AN fibre to a single biphasic current pulse is useful because it avoids the complicating effects of inter-pulse interactions which occur in response to pulse trains. Inter-pulse effects will be considered in Chapter 7, where the model of single-pulse response is used as the basis for a model of pulse-train response.

In preparation for considering the response to a train of current pulses, any single biphasic current pulse  $n$  is considered. A *single* pulse is defined as any pulse whose response properties are not dependent on any previous pulses. The model will therefore be valid for any single pulse presented in isolation, for the first pulse in a pulse train, and for all pulses in a pulse train of sufficiently low pulse rate.

### 5.3.1 Deterministic model of single-pulse response

The basis of the single-fibre model developed in this section is the Hill (1936) threshold model of action potential generation. Hill's model is a phenomenological model that does not directly model any actual physical potentials, conductances or movement of ions. Instead, conceptual 'potentials' are used to describe how a neuron responds to electrical stimulation, determining if and when an action potential will result in response to a particular stimulus.

Hill's model utilises two such 'potentials', which he labels the 'stimulus potential' ( $V$ ) and the 'threshold potential' ( $U$ ). A slightly different nomenclature from Hill is applied when using his equations in order to make clear to which component of the model each symbol refers.  $V_{\text{stim}}$  is used to refer to the stimulus potential and  $V_{\text{thr}}$  to refer to the threshold potential.  $V_{\text{stim}}$  responds to the injection of electrical current with a process Hill calls 'excitation'.  $V_{\text{thr}}$  is also affected by the injection of current in a manner termed 'accommodation'. When the stimulus potential  $V_{\text{stim}}$  crosses the threshold potential  $V_{\text{thr}}$  an action potential is generated.

Hill investigated the behaviour of  $V_{\text{stim}}$  and  $V_{\text{thr}}$  for a range of stimulus waveforms, including monophasic current pulses. He showed how an action potential could result from the change in potential near the cathode in response to rapid increase in current at the onset of the pulse. Additionally, an action potential may result from the change in potential near the anode in response to the current returning to its resting value at the offset of the pulse.

For biphasic current pulses, discharges may occur in response to the onset and offset of both the cathodic and the anodic phases. However, the thresholds are lower for the onset of the cathodic phase (Ranck, 1975; Hill, 1936) and therefore the response to biphasic pulses is approximated by the response to the onset of the cathodic phases. Furthermore, inner hair cell loss is assumed to be complete, such that no spontaneous activity exists (Kiang et al., 1970).

**Approximation 5.3.1** It is considered that discharges may only occur in response to the cathodic phase of a biphasic current pulse and may not occur between pulses. A discharge in response to the cathodic phase of any biphasic pulse  $n$  is referred to as having *occurred during pulse  $n$* .

Hill described the response of the stimulus potential  $V_{\text{stim}}$  to the onset of the cathodic phase according to the equation:

$$V_{\text{stim}}(t) = V_{\text{stim}}(0) + c\tau I \left(1 - e^{-t/\tau}\right)$$

where  $c$  is a constant (Hill uses  $b$ ),  $I$  is the amplitude of the current pulse and  $\tau$  is the time-constant of the change in stimulus potential (Hill uses  $k$ ).

Hill estimated that the threshold potential changes in response to the cathodic phase according to the equation:

$$V_{\text{thr}}(t) = V_{\text{thr}}(0) + c\tau I \left[ 1 + \frac{e^{-t/\tau}}{\lambda/\tau - 1} - \frac{e^{-t/\lambda}}{1 - \tau/\lambda} \right]$$

where  $\lambda$  is the time-constant of the change in threshold potential.

With the correct values of  $c$ ,  $\tau$  and  $\lambda$ , Hill's model predicts whether or not an action potential will occur, and also its latency. In order to increase the ease and efficiency with which it can be calculated whether or not an action potential occurs, a series of further approximations are made. As a consequence of these approximations, the model's prediction of the exact latency of the response will be degraded.

Both  $\tau$  and  $\lambda$  are properties of the neuron, not of the stimulating current, and  $\lambda/\tau$  can typically be on the order of 40–80 for unmyelinated nerve (Hill, 1936). Threshold versus pulse duration curves for myelinated fibres (Bostock, 1983) suggest that they have an even greater ratio. If the stimulus intensity required to generate an action potential is of primary concern, and the exact latency of the action potential is of secondary importance, the ratio of the time-constants is large enough that accommodation can be ignored (Hill, 1936). As a consequence of this approximation, an action potential will occur if the asymptotic value of  $V_{\text{stim}}(t)$  is greater than or equal to  $V_{\text{thr}}(0)$ . The time course of  $V_{\text{stim}}(t)$  can therefore be approximated by a step-function.

**Approximation 5.3.2**  $V_{\text{thr}}$  is considered not to change in response to the stimulating pulse, remaining at a constant value,  $V_{\text{thr}}(n)$ , for the duration of the cathodic phase of pulse  $n$ .  $V_{\text{stim}}$  is considered to increase instantaneously in response to a current pulse and remain at a constant value,  $V_{\text{stim}}(n)$ , for the duration of the cathodic phase of pulse  $n$ .

In making this approximation, a systematic bias is introduced to the estimate of the stimulus intensity required to produce an action potential. The error is equal to  $(\lambda/\tau)^{\frac{1}{\lambda/\tau-1}} - 1$  (Hill, 1936). For a ratio  $\lambda/\tau = 40$  the error will be 9.9%, and increasing the ratio to 80 causes the error to fall to 5.7%. For myelinated nerve, such as the mammalian auditory nerve, the error will be even smaller.

Following Approximation 5.3.1 and 5.3.2, Hill's model can be represented by the electrical circuit diagram shown in Figure 5-2.

This model is in effect a Bernoulli process (trial) (Devore, 1987), where a discharge in response to a pulse is considered to have a value of 1 and no discharge a value of 0. The output of this model is defined as a Bernoulli random variable  $p(n)$  which is dependent on the parameters of pulse  $n$  and the parameters of the neural model.

$$p(n) \triangleq \Pr \{ \text{discharge during pulse } n \}$$

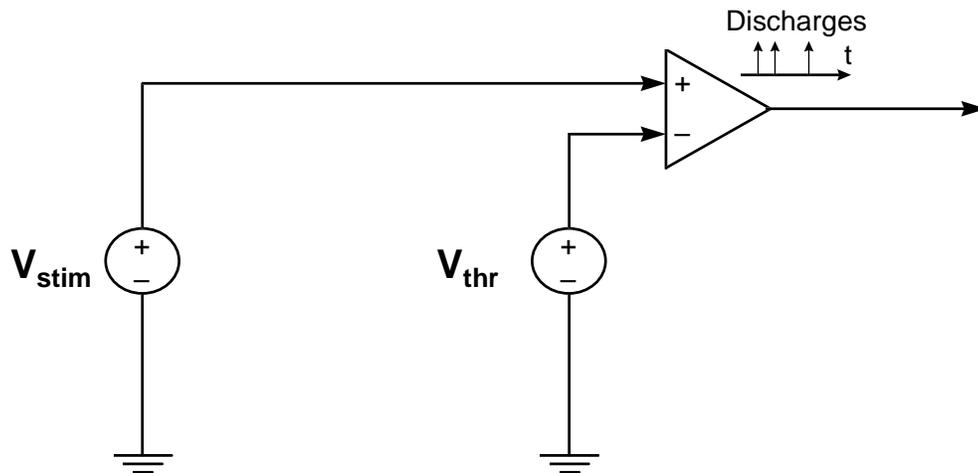


Figure 5-2: Deterministic model of single-pulse response.

An expression for the probability of discharge in response to any single pulse  $n$  can be derived directly from the electrical circuit. For the deterministic model, only two possible values of  $p(n)$  exist, zero and one. If the stimulus intensity is great enough to cause  $V_{\text{stim}}(n)$  to rise to  $V_{\text{thr}}(n)$  or above, an action potential will always result. If the stimulus intensity is less than this critical value, no action potential will ever result. That is:

$$p(n) = \begin{cases} 0 & \text{for } V_{\text{stim}}(n) < V_{\text{thr}}(n) \\ 1 & \text{for } V_{\text{stim}}(n) \geq V_{\text{thr}}(n) \end{cases} \quad (5.1)$$

The probability of discharge versus stimulus intensity relationship can be referred to as an input/output (I/O) function.

### 5.3.2 Prediction of single-pulse response by deterministic model

Figure 5-3 shows the deterministic model fit to cat data from (Javel et al., 1987). The horizontal scale is chosen to correspond to typical behavioural dynamic ranges of cochlear implant users in response to a single pulse (White, 1984b).

The step-function prediction is not due to the simplicity of Hill's model—even the most complex deterministic models are still only able to provide a step-function fit to these data!

The step-function fit is a relatively poor representation of the data, and the practical implications of fitting a step-function are significant. Information which is encoded in the probability of discharge will not be accurately described by the step function. For example, the step-function can only code intensity levels as being above or below some 'threshold' value, whereas the physiological data indicate that the probability of discharge in a single AN fibre could be used to code intensity

levels over some 3 to 4 dB range. Typical behavioural dynamic ranges in cochlear implant users in response to single short-duration pulses are 6 to 10 dB (White, 1984b), and the fibre's actual I/O function covers approximately half of that range. A model which produces a much more accurate fit to these data is therefore required.

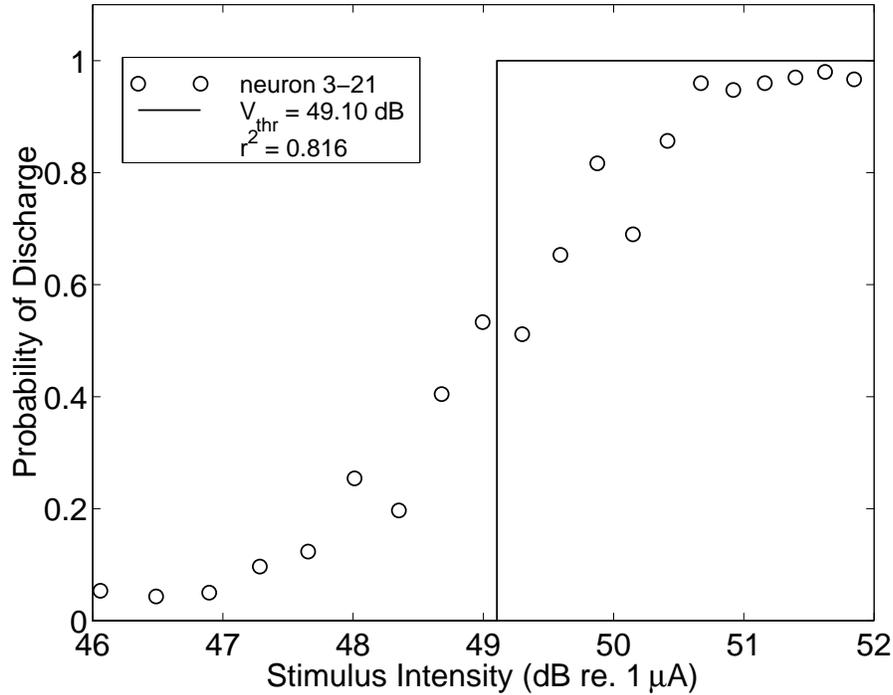


Figure 5-3: Deterministic model (solid line) fit to probability of discharge data (o) from Neuron 3-21 of Javel et al. (1987) for a single biphasic pulse of duration 100  $\mu$ s/phase (*cf.* stochastic model fit to this data plotted in Figure 5-5).

### 5.3.3 Stochastic model of single-pulse response

In this section the model shown in Figure 5-2 is extended to create a stochastic model of single-pulse response. From Verveen's measurement of membrane noise it may seem intuitive to add a noise source to the stimulus potential  $V_{\text{stim}}$ . However, comparison of Hill's model with Hodgkin-Huxley type models indicates that  $V_{\text{thr}}$  is an approximation of the behaviour of the ionic channels which are the source of the noise, and is therefore also an appropriate choice for including the noise. In fact, subtracting a noise potential from  $V_{\text{thr}}$  is computationally identical to adding it to  $V_{\text{stim}}$ . Furthermore, the mean of the noise is zero and it has a symmetrical amplitude distribution, so it makes no difference whether the noise is added or subtracted. In this thesis it is chosen to add the noise potential  $V_{\text{noise}}$ , which has a  $1/f$  spectrum and a Gaussian distribution, to  $V_{\text{thr}}$ , producing the model shown in Figure 5-4.

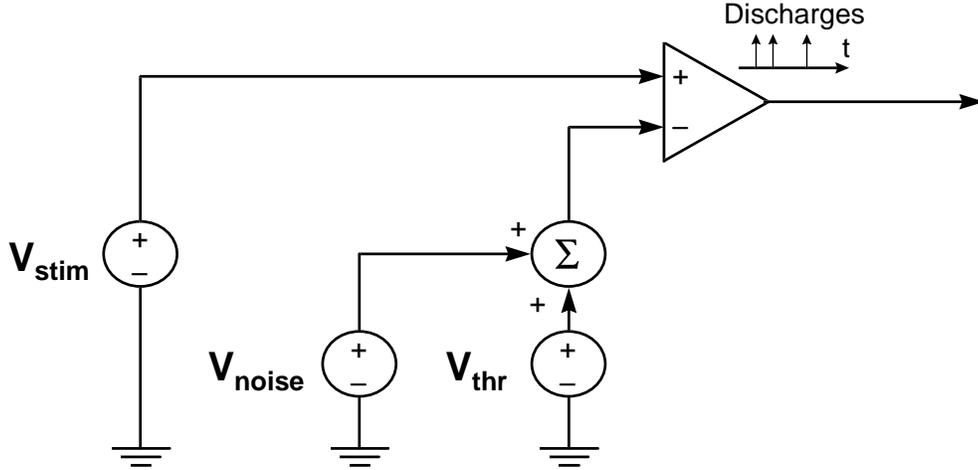


Figure 5-4: Stochastic model of single-pulse response.

Accurate simulation of noise with a  $1/f$  frequency spectrum is quite laborious. However, the  $1/f$  shape of the spectrum is such that the low-frequency content of the noise outweighs the high-frequency content to such an extent that  $V_{\text{noise}}$  changes slowly enough that it may be considered to remain constant for the duration of a short current pulse.

**Approximation 5.3.3**  $V_{\text{noise}}$  is considered to have a constant value  $V_{\text{noise}}(n)$  for the duration of pulse  $n$ . This noise is normally distributed with a mean of zero and variance  $\sigma^2$ .

**Lemma 5.3.1** Given the model of action potential generation described by Figure 5-4 and Approximations 5.3.1, 5.3.2 and 5.3.3, the probability of discharge in response to a single current pulse can be approximated by the equation:

$$p(n) = \frac{1}{2} \left( 1 + \operatorname{erf} \left( \frac{V_{\text{stim}}(n) - V_{\text{thr}}(n)}{\sqrt{2}\sigma} \right) \right) \quad (5.2)$$

**Proof:** see Appendix B.

The function  $p(n)$  is dependent on  $V_{\text{stim}}(n)$  with two independent variables:  $V_{\text{thr}}(n)$  and  $\sigma$ . As with the deterministic model,  $V_{\text{thr}}(n)$  is referred to as the ‘threshold’. From (5.2), this definition of threshold corresponds to the stimulus level which produces a probability of firing of 0.5 and is consistent with Verveen’s convention for defining threshold.

Verveen referred to the second variable  $\sigma$  in a normalised form, which he labelled the ‘Relative Spread’ (RS) (Verveen, 1961; Verveen, 1960):

$$\text{RS} \triangleq \frac{\sigma}{\text{threshold}} \quad (5.3)$$

The greater the value of RS, the shallower the slope of the I/O function.

### 5.3.4 Prediction of single-pulse response by stochastic model

The integrated-Gaussian (error function) description of single-pulse probability of discharge given by (5.2) is consistent with both physiological results (Dynes, 1996; Hoopen and Verveen, 1963; Verveen, 1961; Verveen, 1960) and modelling studies (Rubinstein, 1995; Lecar and Nossal, 1971) investigating the response of nerve to monophasic pulses. However, it is useful to investigate how accurately this description predicts data from the cat AN in response to biphasic pulses.

Figure 5-5 shows the stochastic model fit to the same cat data as plotted in Figure 5-3. The stochastic model predicts the neuron's I/O function much better than the deterministic version does, with a correlation coefficient of 0.99. Consequently, this model will better describe information encoded in the probability of discharge.

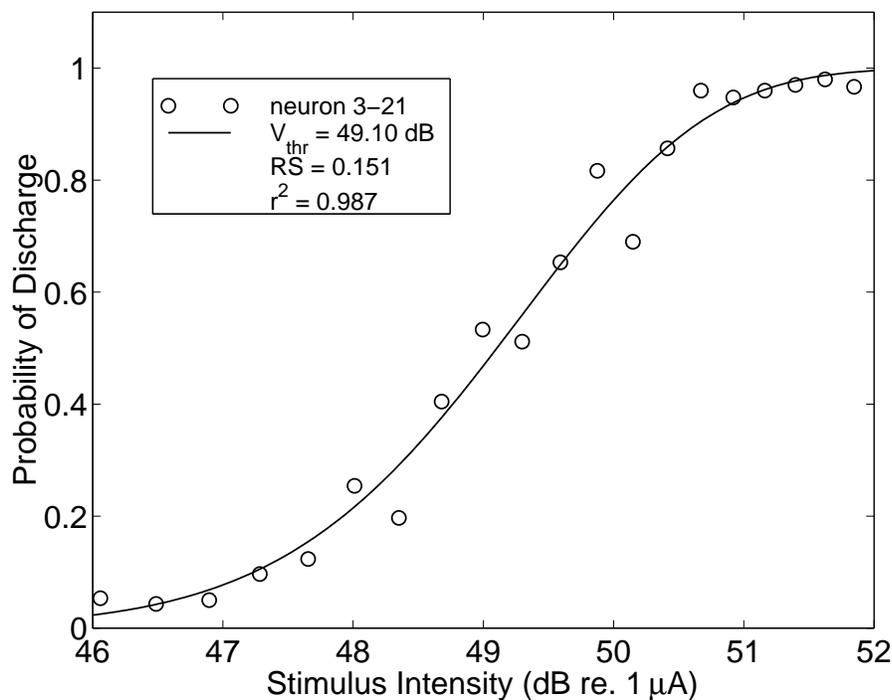


Figure 5-5: Stochastic model (solid line) fit to probability of discharge data (o) from Neuron 3-21 of Javel et al. (1987) for a single biphasic pulse of duration 100  $\mu$ s/phase (*cf.* deterministic model fit to this data plotted in Figure 5-3).

Further examples of integrated-Gaussian function fits to cat data from Javel et al. (1987) are given in Figures 5-6 and 5-7 for single biphasic pulses of durations 200  $\mu$ s/phase and 400  $\mu$ s/phase respectively. The quality of these fits suggests that the assumptions and approximations which have been made thus far in the development of the model have been reasonable and accurate, at

least for the range of stimulus parameters investigated in this chapter.

## 5.4 Total auditory nerve response

In order to model the response of a population of auditory nerve fibres, it is necessary to determine the model parameters for each neuron in the population, as well as the intensity of the excitatory current at the initial site of action potential generation in each neuron.

To demonstrate how this may be done, a population model for the total auditory nerve is developed following the approach used by O’Leary et al. (1995), but using the model of single-fibre response developed in this chapter. This is achieved by fitting the model parameters to the AN statistics from cat data and by using a simple model of current spread.

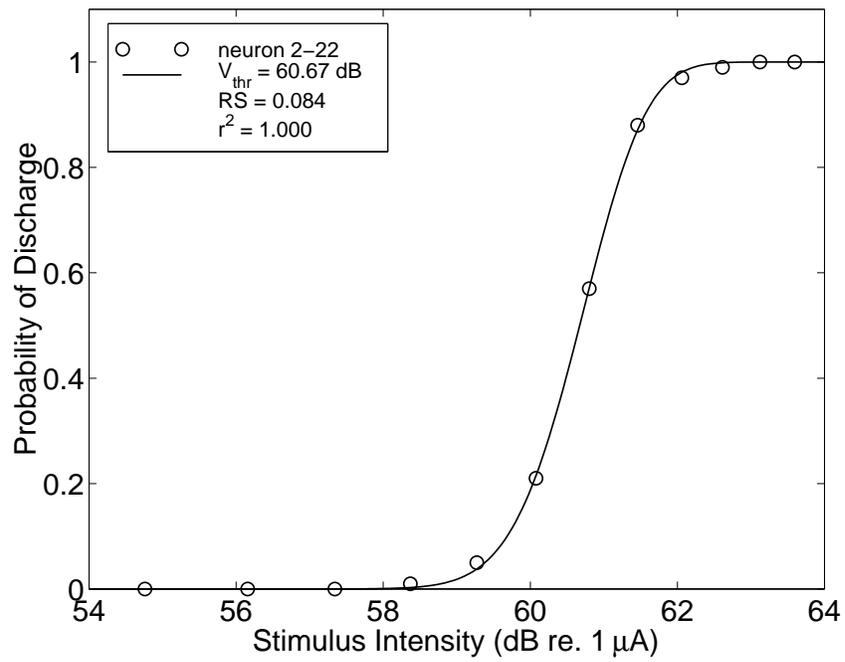
### 5.4.1 Fitting of model parameters to AN statistics from cat

Javel et al.’s (1987) data from single AN fibres in the cat have been analysed, in order to estimate the distribution of single-pulse I/O functions present in the cat AN. Stochastic model fits to 15 neurons had a mean correlation coefficient of 0.99, with a minimum of 0.94.

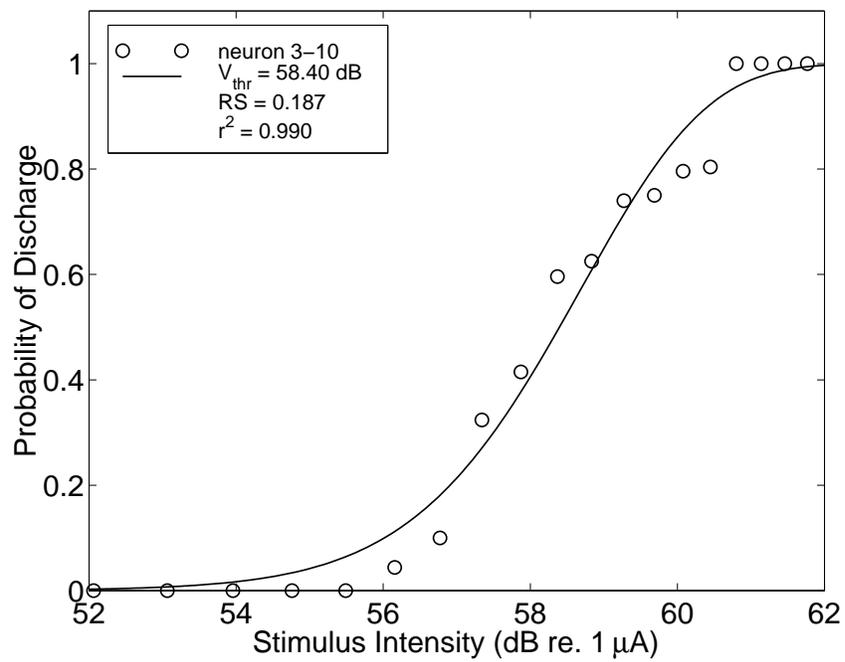
In Figure 5-8 relative spread is plotted against threshold for each neuron in response to a single biphasic pulse of duration 200  $\mu\text{s}$ /phase or 400  $\mu\text{s}$ /phase. The RS values at a pulse duration of 200  $\mu\text{s}$ /phase have a mean of 0.12 and a standard deviation of 0.05, and a pulse width of 400  $\mu\text{s}$ /phase yields similar values. This range of RS values is consistent with other data from the cat AN (Miller et al., 1997).

The distribution of threshold values is confounded by the fact that the distance from the stimulating electrode to the fibre, or more exactly the distance to the initial site of action potential generation in the fibre, is unknown. van den Honert and Stypulkowski (1987a) have plotted AN fibre thresholds to intracochlear electrical stimulation against cochlear place (as inferred from the AN fibre’s characteristic frequency to acoustical stimulation). Since the cochlear place of the stimulating electrodes were known, it was therefore possible to estimate the site of each AN fibre relative to the stimulating electrode. For monopolar and longitudinal bipolar stimulation, the range of thresholds for fibres close to the stimulating electrode was on the order of 10–14 dB. This is consistent with the range of thresholds seen in Javel et al.’s (1987) cat data, considering the effect of distance from the stimulating electrode, and with other data from the cat AN (Miller et al., 1997).

It has been suggested (see Chapter 6) that changes in the relative spread following electrical pulses of different durations might be involved in the sharp decrease of human psychophysical thresholds to pulsatile electrical stimulation (Pfungst et al., 1991; Shannon, 1989). To extend the model such that it can be used to test this hypothesis, the RS was determined for neural responses

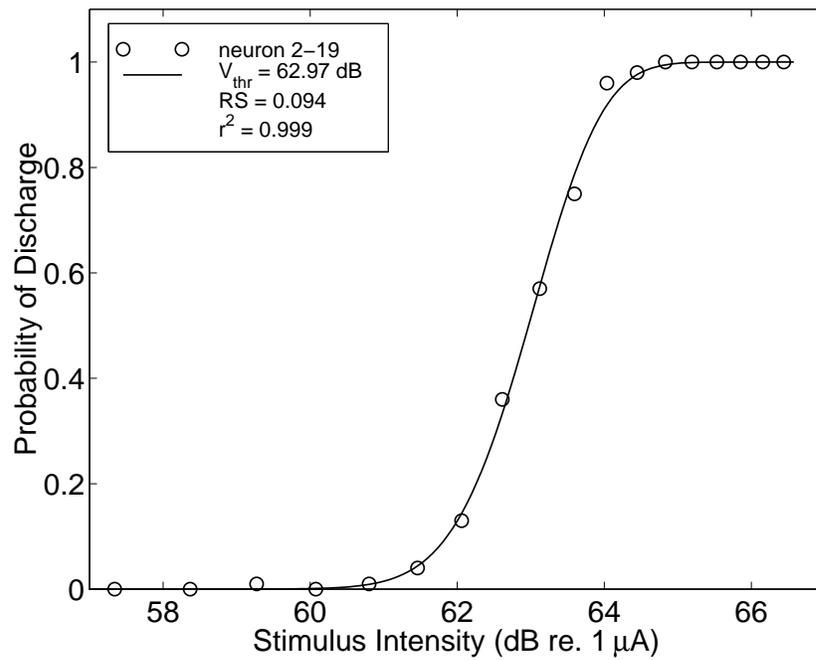


(a)

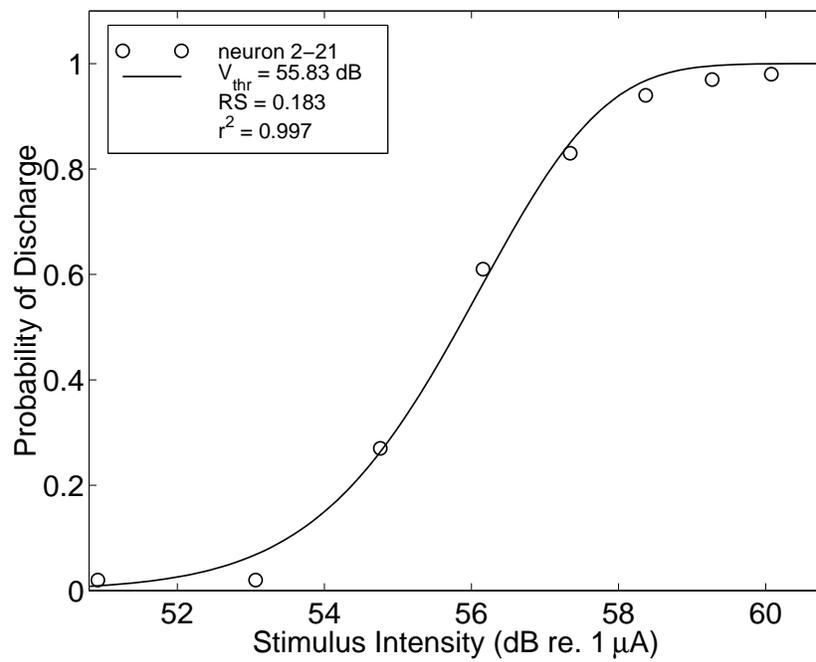


(b)

Figure 5-6: Fits to fibre I/O functions from Javel et al. (1987) for a single pulse of duration 200  $\mu$ s/phase.



(a)



(b)

Figure 5-7: Fits to fibre I/O functions from Javel et al. (1987) for a single pulse of duration  $400 \mu$ s/phase.

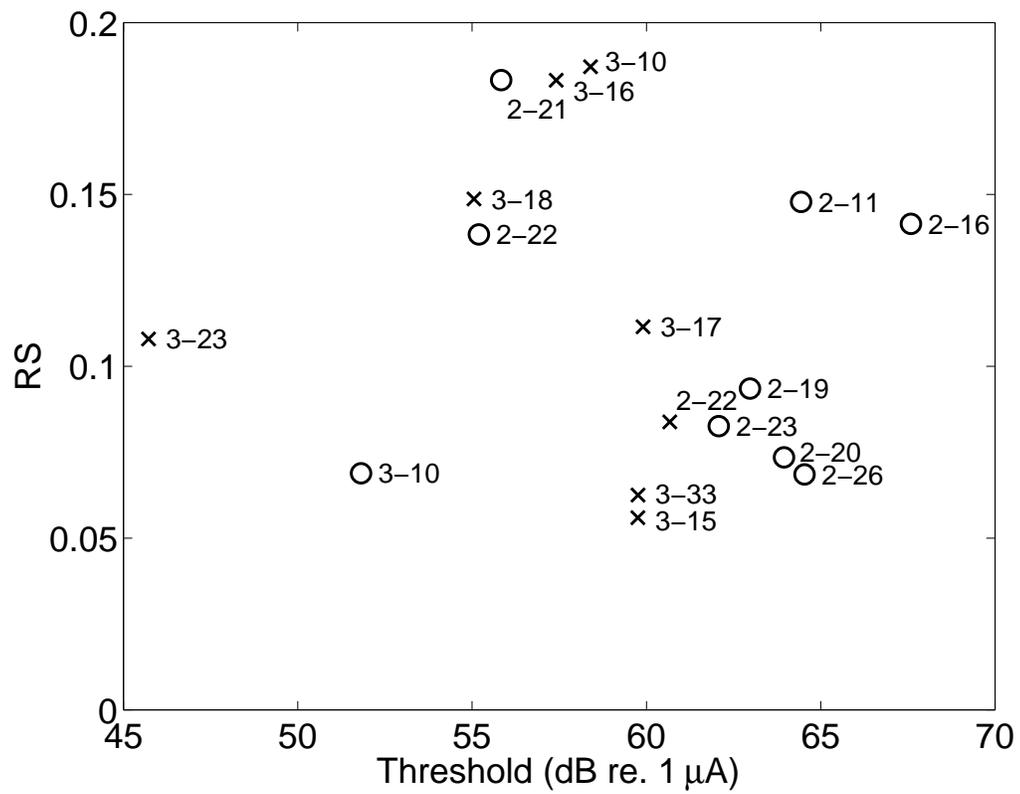


Figure 5-8: Relative spread versus threshold for neurons from Javel et al. (1987) as labelled, in response to a single biphasic pulse of width 200  $\mu$ s/phase (x) or 400  $\mu$ s/phase (o).

to symmetric anodic/cathodic biphasic stimuli of differing durations.

Figure 5-9 shows discharge probability versus level curves for three different phase durations in a single unit. As the phase duration increases, the slope of the curve becomes shallower, indicating a greater dynamic range. Computing the RS from these curves shows that the RS increases as the phase duration of the anodic/cathodic biphasic stimulus increases.

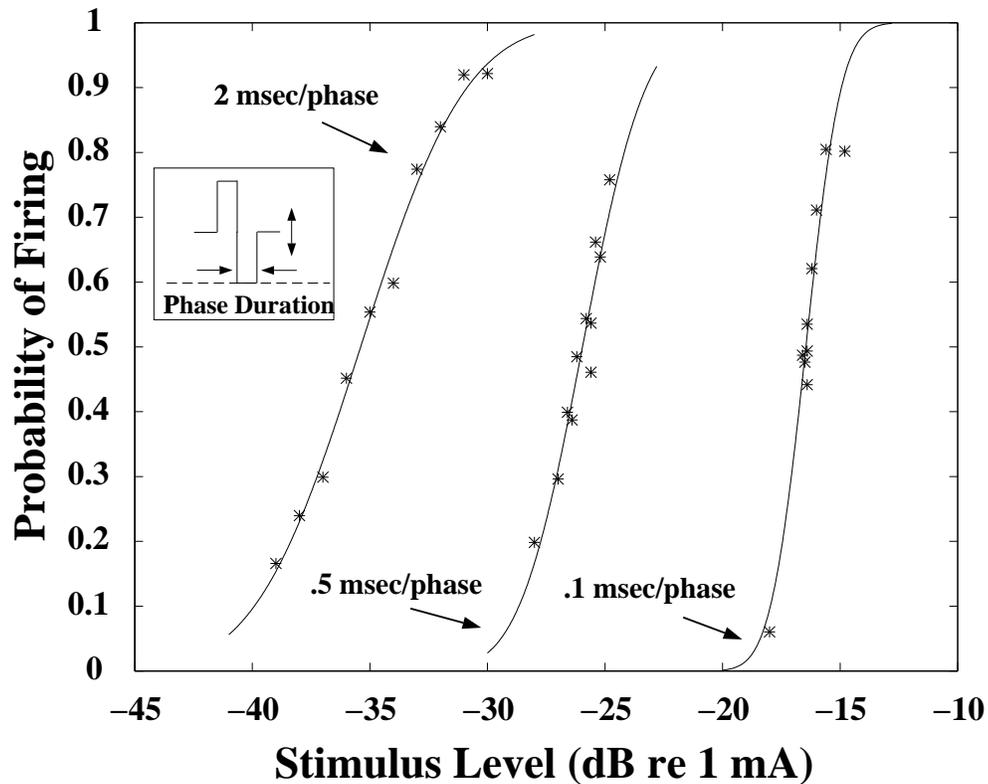


Figure 5-9: Discharge probability versus level functions for single symmetric biphasic anodic/cathodic pulses of differing durations. For each duration, several measures of the probability of discharge were made at different stimulus levels. Individual measures consisted of 100 stimulus presentations and are represented by asterisks. Curves are error function fits determined by a least-squares method. All data are from a single unit.

Figure 5-10 makes this clear for three units from the same cat. For every unit, the RS increases as the duration per phase of the stimulus increases. The increase of RS with increasing pulse width as seen in this figure has also been observed in cat data from a different study (Smith, 1996). Even earlier evidence of the potential importance of phase duration on auditory input/output slopes was observed in the anteroventral cochlear nucleus using sinusoidal stimuli (e.g., see Figure 33 of White, 1984b).

Integrated-Gaussian fits to the complete data set are shown in Figure 5-11. The curves were generated from (5.2) using the mean values of threshold and RS at each pulse width. Even at some

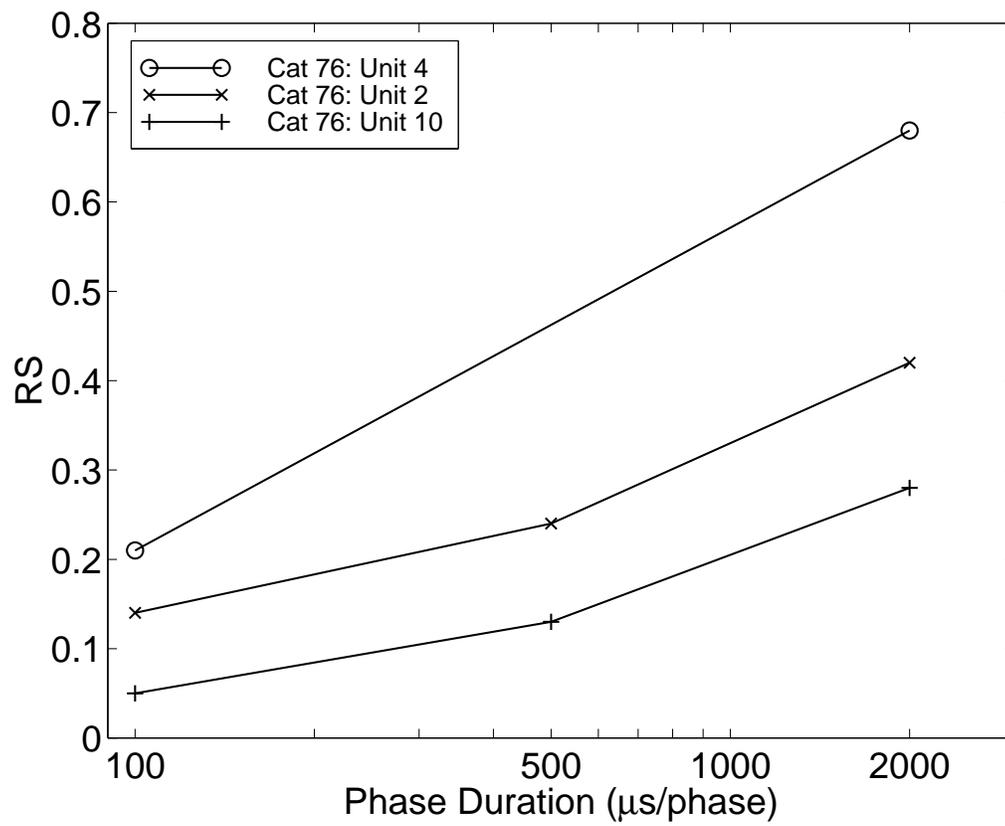


Figure 5-10: Relative spread as a function of phase duration for three units from the same cat. The relative spread was determined as described in the text; individual measures are represented by symbols.

of the lower pulse widths, there are significant changes in the slopes of the discharge probability curves. For example, the slope changes by approximately 1.67:1 when the pulse width is lowered from 2 000 to 500  $\mu\text{s}/\text{phase}$ .

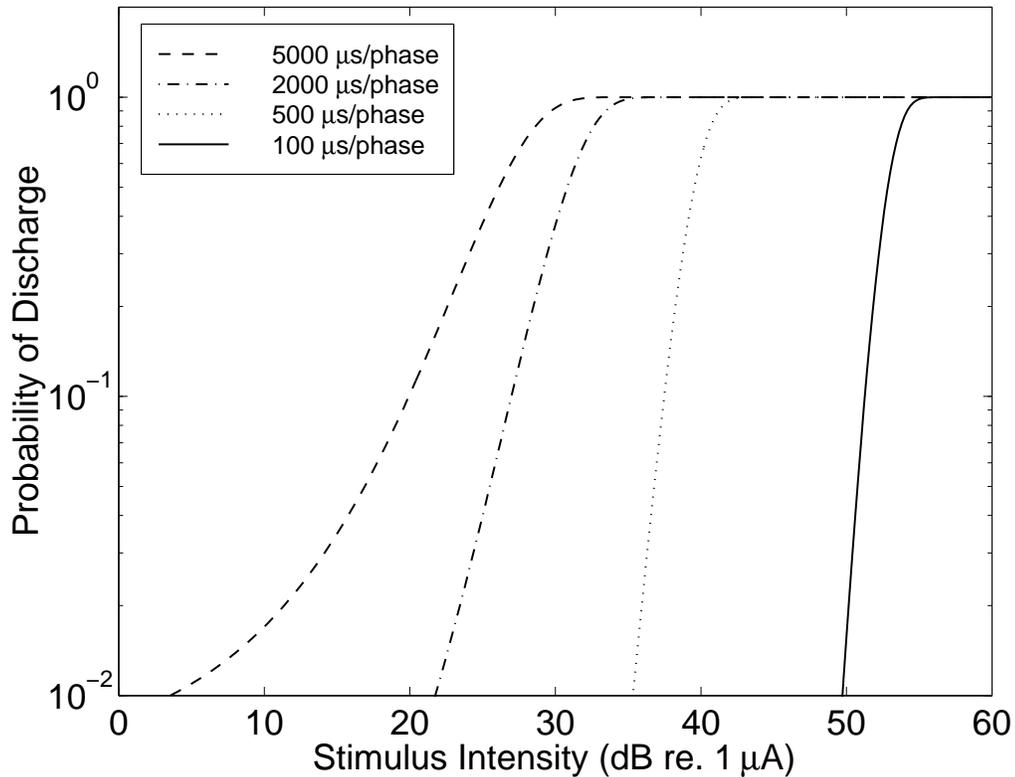


Figure 5-11: Fits to probability of discharge versus stimulus intensity data for pulse widths of 100  $\mu\text{s}/\text{phase}$  to 5 000  $\mu\text{s}/\text{phase}$ .

There are insufficient data to examine how the distributions of threshold and RS covary with pulse width. Therefore it is assumed that the distributions do not change with pulse width. It may also be feasible that the standard deviation of RS covaries linearly with its mean—further data would be required to determine this.

The available data (see Figure 5-8) do not support the hypothesis that RS is a function of threshold. When more data become available it may be possible to detect such a dependency. Meanwhile, it is assumed that the distribution of single-pulse I/O functions present in the normal cat cochlea can be well described by two independent random variables, ‘Threshold’ and ‘RS’. The means of Threshold and RS as functions of pulse width are estimated by fitting appropriate functions<sup>3</sup> to the values of threshold and RS of the curves in Figure 5-11, producing the relationships:

<sup>3</sup>Particular care should be taken in extrapolating (5.5) to points beyond 5 000  $\mu\text{s}/\text{phase}$ , because the function becomes nonmonotonic at pulse widths greater than 6 000  $\mu\text{s}/\text{phase}$ , but there is no evidence to suggest that this effect actually occurs in neurons.

$$E[\text{Threshold}] = 121.04 \times PW^{-0.18} \text{ (dB re. } 1 \mu\text{A)} \quad (5.4)$$

and

$$E[\text{RS}] = 0.12 + 9.51 \times 10^{-5} \times PW - 7.90 \times 10^{-9} \times PW^2 \quad (5.5)$$

where  $PW$  is the pulse width (phase duration) in  $\mu\text{s}/\text{phase}$ .

Taking a conservative estimate of the distribution of thresholds seen in the cat data, threshold is taken to be uniformly distributed  $\pm 5$  dB around its mean as given by (5.4). From Figure 5-8, the standard deviation of RS is approximately half its mean at short pulse widths (200 and 400  $\mu\text{s}/\text{phase}$ ). Therefore RS is taken to be normally distributed with a standard deviation of 0.06 (half the mean of RS as given by (5.5) at short pulse widths).

### 5.4.2 Current spread

Following O'Leary et al. (1995), the electrode array is approximated by a point source of current at the active electrode and the auditory nerve tissue by a homogeneous resistive medium consisting of a uniform density of single AN fibres. The potential at the site of each auditory neuron can be calculated for a given stimulus level. From O'Leary et al.'s (1985) data, the current decays at a rate of 3–6 dB/mm for bipolar stimulation. This model is extended to also give the current spread for monopolar stimulation. The attenuation is taken to be approximately 0.5 dB/mm for monopolar stimulation and 4 dB/mm for bipolar stimulation<sup>4</sup>. These attenuation rates were obtained from neurophysiological experiments (Merzenich and White, 1977) that measured the spread of excitation across the cochlea as a function of stimulus intensity and electrode configuration. Modelling an electrode placed 15 mm inside a cochlea 30 mm long produces attenuation curves as shown in Figure 5-12.

### 5.4.3 Output of total auditory nerve model

Here the output is presented of a model of 10 000 neurons<sup>5</sup> spread uniformly across the auditory nerve, assuming the statistics of model parameters given in Section 5.4.1 and the model of current spread given in Figure 5-12. In Figure 5-13 the number of discharges summed across all fibres in response to a single biphasic pulse is plotted. For stimulus intensities above approximately 50 dB, the deterministic model and the stochastic model produce very similar predictions of the mean number

---

<sup>4</sup>Bipolar stimulation produces a greater rate of attenuation than monopolar stimulation, because the return electrode is closer to the active electrode, creating a tighter current path. Note that these linear functions provide only a rough approximation of the electrode array, current path and cochlear geometries.

<sup>5</sup>This is a conservative estimate of the number of surviving auditory nerve fibres in a subject with a profound sensorineural hearing loss (Hinojosa and Marion, 1983).

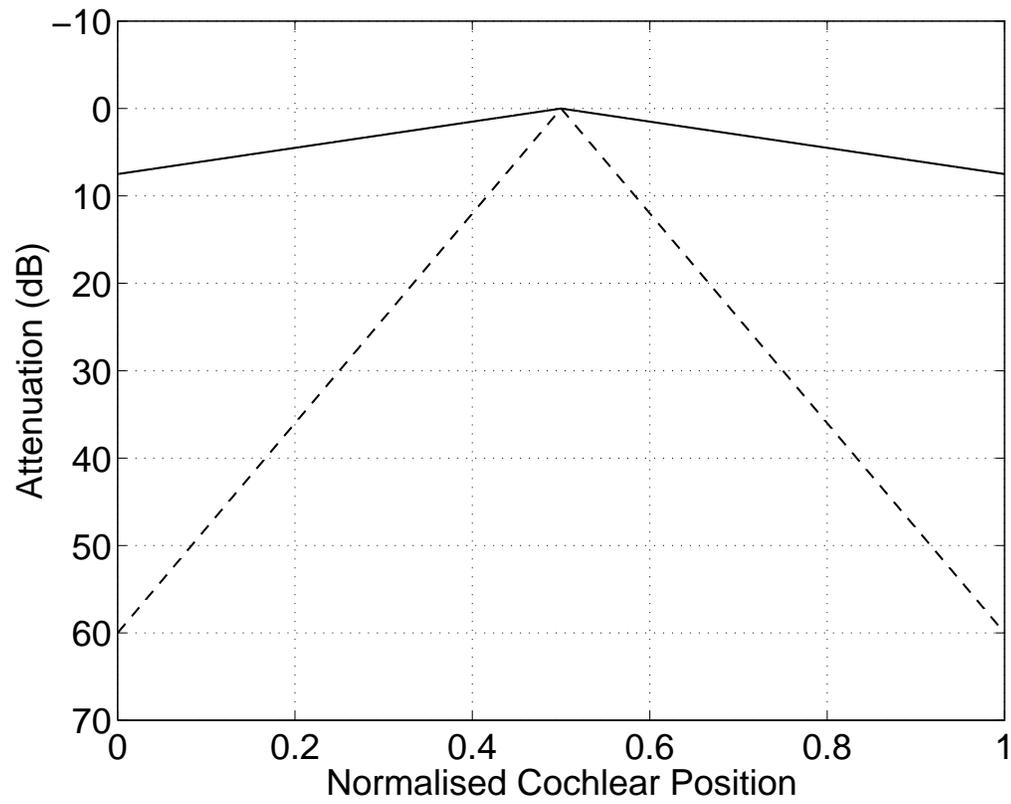


Figure 5-12: Attenuation of the stimulus across the cochlea for monopolar (solid line) and bipolar (dashed line) electrode configurations.

of discharges. However, below 50 dB the stochastic model predicts a quite different response from the deterministic model's. Note that the stochastic model predicts a mean number of discharges of greater than 70 spikes in response to a single biphasic pulse of intensity 50 dB in bipolar mode and greater than 630 spikes in monopolar mode. It is very likely that auditory threshold is reached with an even smaller number of responses, so it is highly feasible that cochlear implants are operating in the region where the stochastic activity is significant (see Chapter 6). Furthermore, the deterministic model will predict zero variance in the number of discharges, whereas the stochastic model will predict non-zero variance (see error bars in Figure 5.13(b)).

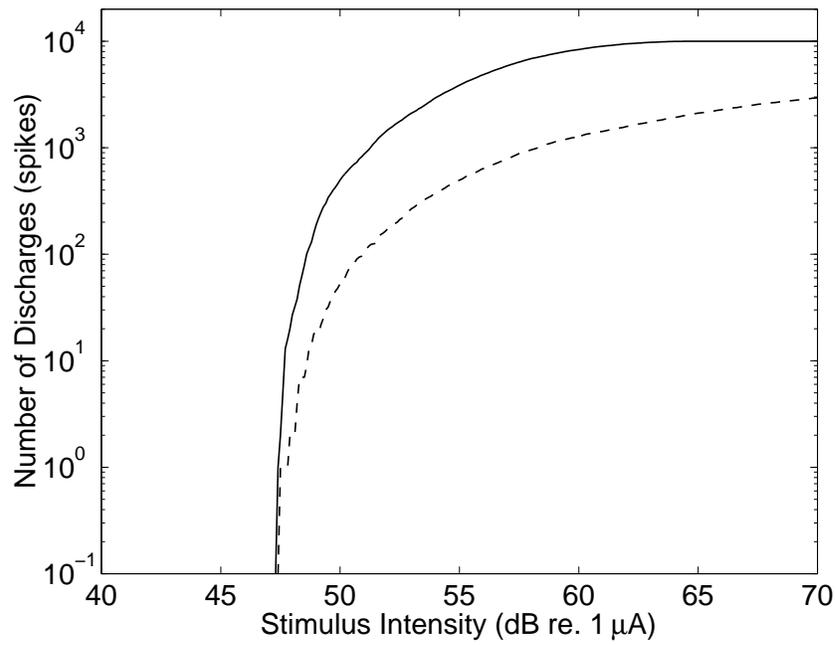
The summing of these Bernoulli random variables is called *Poissonian binomial sampling* (Johnson et al., 1992), which is similar to a binomial process, except that the probability of success (probability of discharge in this case) is different, in general, for each neuron. When the number of neurons is large, the Poissonian binomial distribution is well approximated by a Poisson distribution if the mean number of responses is less than 15 and by a Gaussian (Normal) distribution for mean counts greater than or equal to 15, although the distribution will only have discrete integer values in both cases.

Figure 5-14 illustrates results from 10 000 iterations of Monte Carlo simulations for a very low amplitude stimulus (Figure 5.14(a)) and a moderate amplitude stimulus (Figure 5.14(b)), with a Poisson distribution fit to the former and a Gaussian distribution fit to the latter. The fits appear to be quite good.

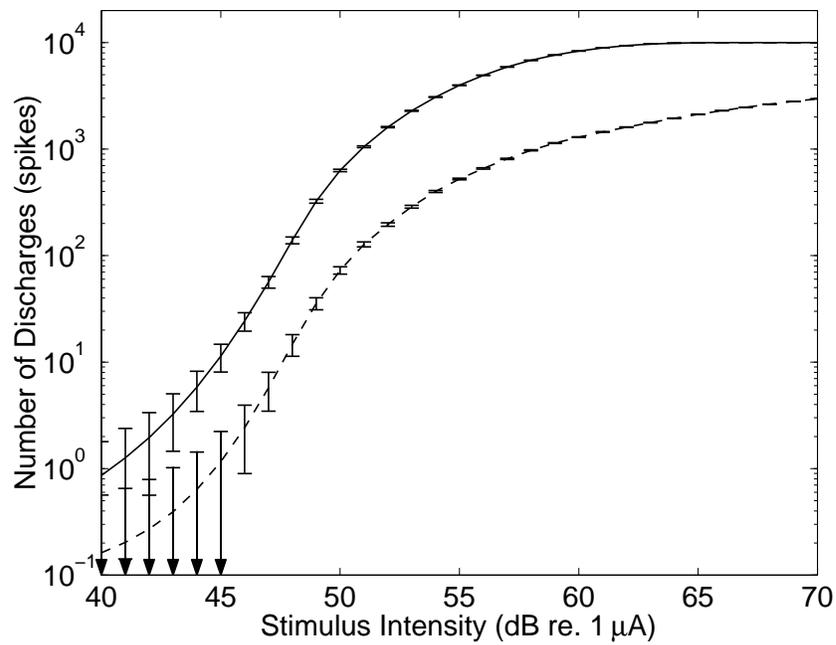
## 5.5 Discussion

The results of this chapter have shown that a deterministic model is a poor description of a neuron's response to single electrical pulses, in light of the improvements in the prediction of physiological data when a stochastic component of response is added to the neural model. Specifically, the results show that a simple stochastic model can accurately predict discharge probabilities present in cat data in response to single biphasic pulses. The results suggest that just two parameters (threshold and RS) can well describe the response of an auditory nerve fibre to single biphasic pulses.

The analytical descriptions of single fibre response to single electrical pulses which developed in this chapter provide a simple and computationally efficient method of modelling the response of a large-scale population of AN fibres to pulsatile electrical stimulation as produced by cochlear implants. Furthermore, the output of the model, being a probability of discharge for each neuron, with each neuron described by a Bernoulli process, allows the direct use of signal detection theory to determine the resolution with which the AN can convey information about an electrical stimulus (see Chapter 6).

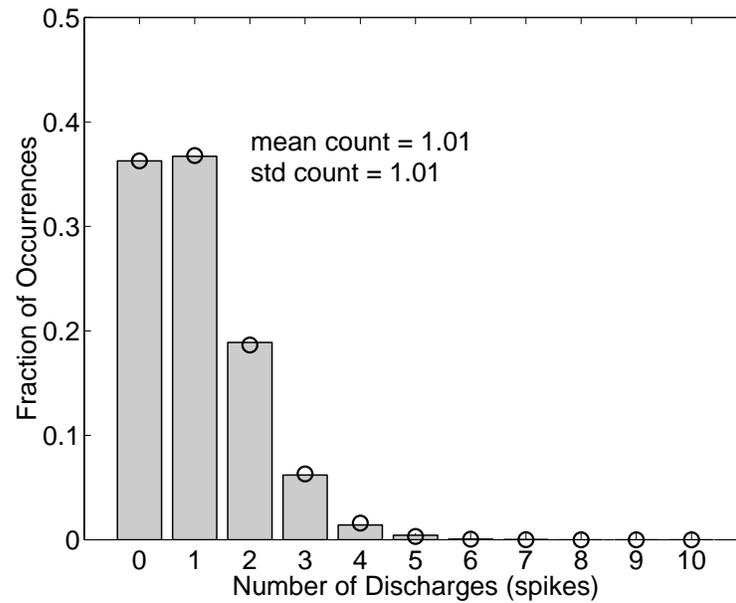


(a) Deterministic Model

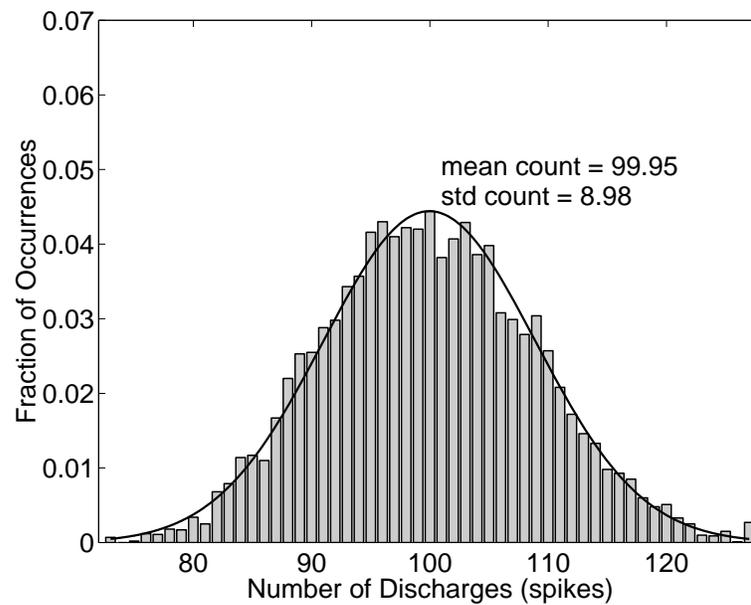


(b) Stochastic Model

Figure 5-13: Number of discharges versus stimulus intensity from model of 10 000 fibres in response to a single biphasic pulse for monopolar (solid line) and bipolar (dashed line) stimulation. Error bars indicate  $\pm 1$  std.



(a) Low-intensity stimulus



(b) Moderate-intensity stimulus

Figure 5-14: Distribution of total number of discharges predicted by 10 000 iterations of Monte Carlo simulations in response to a single biphasic pulse for two different intensities, with a Poisson distribution fit (o) for the case where the mean count less than 15 and a Gaussian (Normal) distribution fit (solid line) for the case where the mean count is greater than 15.

Further physiological and modelling studies would be helpful in refining and extending a number of aspects of the model.

In this chapter randomness in the discharge probability in response to electrical stimuli has been investigated. Physiological studies have shown that randomness occurs not only in the probability of an action potential being generated, but also in its latency (Javel et al., 1987; van den Honert and Stypulkowski, 1987b; van den Honert and Stypulkowski, 1984; Verveen and Derksen, 1968; Verveen and Derksen, 1965). The term ‘jitter’ is used qualitatively to describe this phenomenon, or quantitatively to refer specifically to the standard deviation of the latency (van den Honert and Stypulkowski, 1984). In developing an analytical description of probability of discharge the model’s description of the timing of the discharges has been degraded. A separate stochastic description of action potential latency would be needed to investigate the effects of jitter in theories of coding which are dependent on the exact timing of discharges.

The approximation that discharges occur only in response to the cathodic phase of the biphasic pulse (Approximation 5.3.1) has not appeared to degrade the model’s fits of the physiological data. However, physiological studies have shown that some AN fibres do exhibit a shift in the phase of the biphasic current pulse to which they respond (van den Honert and Stypulkowski, 1987b). Such a phenomenon may have an effect on the I/O function of these fibres. For example, the cause of the increase in RS with increasing pulse width is as yet unknown—in this chapter only a phenomenological description of this effect is provided. This behaviour is absent in responses to monophasic pulsatile stimuli (Dynes, 1996; Rubinstein, 1995; Lecar and Nossal, 1971), suggesting that the effect is in fact due to the addition of the second phase in the stimulus. There are a number of ways in which the second phase could increase RS. One possibility is that the relative level of noise increases with pulse width. Figure 5.1(b) shows that the membrane noise tends to increase when the transmembrane potential deviates from its resting potential. This means that the first phase of the pulse could change the membrane potential and thereby increase the relative noise level. If the membrane potential changes more for longer pulse widths, it may explain the increase in RS. A second possibility is that the initial site of activation changes with alterations in pulse width. Different activation sites (i.e., nodes of Ranvier) in the one fibre could have significantly different relative spreads.

The fits given in Figure 5-11 were to data covering a range of probability of discharge from approximately 0.05 to 0.95. These fits may be invalid for discharge probabilities less than 0.05 at long pulse widths. Figure 5-15 shows that the probability of discharge asymptotes to a non-zero value as the stimulus intensity approaches zero and that this value increases with pulse width. This is not physiologically feasible however, since the pulse width of a non-existent (zero intensity) current pulse cannot affect the probability of discharge. Consequently, it can be concluded that the integrated-Gaussian fit is not valid for very low probabilities of discharge. Further data need

to be collected to show how the probability of discharge changes with pulse width for probabilities considerably smaller than 0.05.

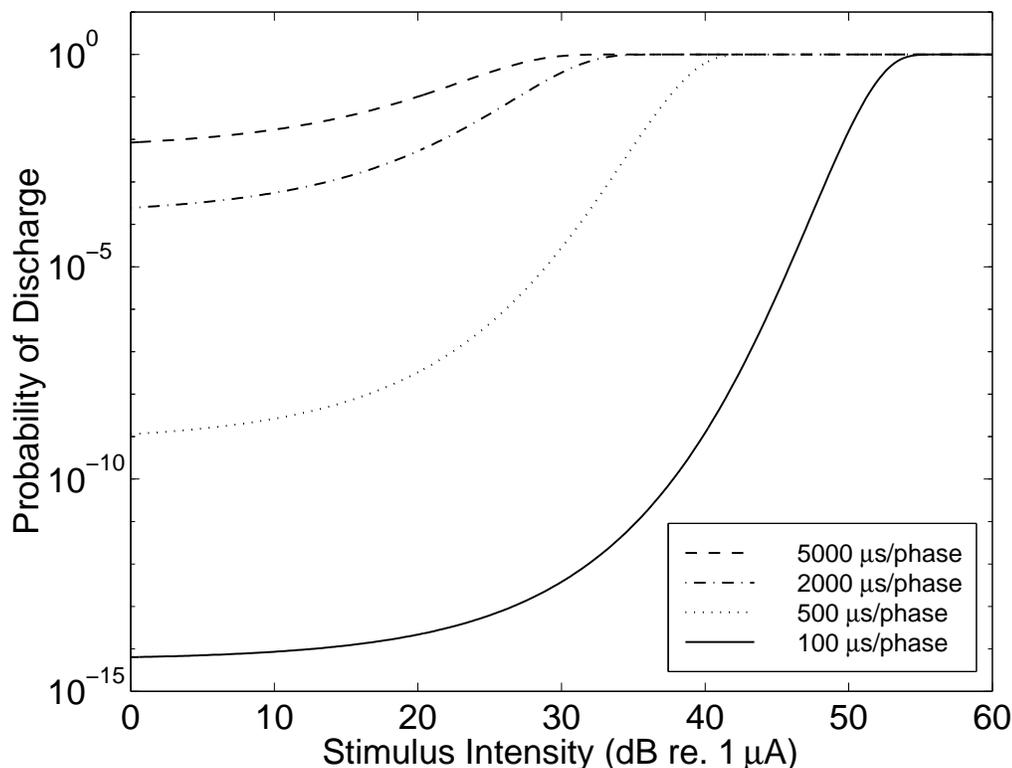


Figure 5-15: Figure 5-11 replotted with an extended ordinate.

The small amount of data used in Section 5.4.1 to investigate how the distributions of threshold and RS vary with stimulus parameters lead to somewhat rough approximations of their behaviour. A study similar to van den Honert and Stypulkowski's (1987a), but in which RS values are collected along with threshold values, would be useful in further refining the statistics of the total AN model. In particular the relationships between site of the electrode, pulse width, threshold and RS need to be further examined.

The physiological data presented in this chapter and the models derived from this data were from cats. Care should be taken in extrapolating the model to other species, including human cochlear implant users. Furthermore, the physiological data are from normal hearing cats. Other than assuming zero inner hair cell survival and allowing an arbitrary number of surviving AN fibres, this model does not take into account the effects which prolonged deafness has on the response of AN fibres to electrical stimulation (Shepherd and Javel, 1997). An extension to this study could be to model the effects of various aetiologies on single-fibre I/O functions.

If the normal behavioural operating range of cochlear implants falls totally in the region above 50 dB in Figure 5-13, then the deterministic model will adequately predict the mean number of

discharges, although it will be unable to predict the variance in the number of discharges. However, if some or all of the operating range falls in the region below 50 dB, Figure 5-13 indicates that the two models predict very different responses. Physiological and behavioural studies in the same animals provide one method of determining the intensities at which the physiological data should be collected (Smith, 1996). An alternative method is to model how an ideal observer would behave from the output of the AN model. Such an evaluation is carried out in the next chapter, where a range of psychophysical measures are predicted using this AN model. The results of this study suggest that cochlear implants are indeed operating in the region where stochastic activity is significant. In all the cases investigated, the stochastic model predicted psychophysical performance significantly better than the deterministic model did. This suggests that the stochastic model is not just more accurate in its prediction of physiological response, but also in its prediction of the resulting behavioural performance.

## Chapter 6

# The effects of stochastic neural activity in a model predicting loudness perception with cochlear implants

### 6.1 Introduction

Stochastic activity in the Auditory Nerve (AN) will affect what discharge patterns exist in response to acoustical or electrical stimulation, how these discharges will be processed and propagated by the various sections of the auditory pathways, and ultimately the perceptual performance of the auditory system. However, models of cochlear implant perception have historically utilised deterministic descriptions of AN response to electrical stimulation, which ignore stochastic activity present in the response. Recently, physiological models of AN response have been developed which do incorporate stochastic activity (Bruce, Irlicht, White, O'Leary, Dynes, Javel and Clark, 1997a; Bruce, Irlicht, White, O'Leary, Dynes, Javel and Clark, 1997b; O'Leary et al., 1998; Irlicht and Clark, 1996; Dynes, 1996; Rubinstein, 1995; White et al., 1987; Hochmair-Desoyer et al., 1984; White, 1978), but the consequences for behavioural performance have not been investigated.

In the normal mammalian auditory system, the AN exhibits spontaneous firing of fibres in the absence of acoustical stimulation and also exhibits variance in the timing of discharges in the presence of acoustical stimulation (Kim et al., 1990; Kiang et al., 1965). The inner-hair-cell/auditory-nerve synapse appears to be the primary source of this stochastic activity (Sewell, 1996). In most cases, profound sensorineural hearing loss is the result of the functional loss of all,

or nearly all, hair cells. As a consequence, hair cells are unlikely to be a significant source of noise in the profoundly deaf (Kiang et al., 1970).

A series of studies by Verveen and Derksen (1968) have shown, however, that random fluctuations in a neuron's membrane potential produce a secondary source of noise which will become more apparent after loss of the inner hair cells. The magnitude of these fluctuations is significantly less than the noise from the inner hair cells and has consequently been regarded as unimportant for some time. Hence, neural response to electrical stimulation has historically been modelled via deterministic descriptions such as the Hodgkin–Huxley equations. However, potentially significant variance has been measured in the response of AN fibres to single current pulses (see Chapter 5) and pulse trains (see Chapter 7), which cannot be explained at all by deterministic models. This stochastic activity will affect electrically elicited AN responses and consequently could be a significant factor in the psychophysical performance of cochlear implant users.

Studies of single fibre response, where only an arbitrary measure of threshold is recorded, do not accurately predict behavioural threshold versus phase duration (strength-duration) curves for sinusoidal stimulation (Pfungst, 1988) or for pulsatile stimulation (Pfungst et al., 1991; Pfungst, 1990). Indeed, strength-duration curves of cochlear implant users are not well predicted by deterministic models of AN response to electrical stimulation (Pfungst, 1990; Shannon, 1989).

In this chapter it is investigated whether these and other inaccuracies in predictions of loudness perception could be due to ignoring the stochastic response of the AN to electrical stimulation. In order to avoid the complication of inter-pulse interactions and to enable the use of a simpler and computationally faster single-fibre model, the investigation in this chapter is restricted to low-rate pulse trains. A model of loudness based on the description of neural excitation developed in Chapter 5 is derived and model predictions for the deterministic and stochastic descriptions of single-fibre response to electrical stimulation are compared. The psychophysical section of the model is developed in such a way that signal detection theory can be applied to directly predict how behavioural threshold, dynamic range and intensity difference limens change with stimulus parameters.

In Section 6.2, Methods, the psychophysical model and the sources of the psychophysical data which appear in the Results section are described. In Section 6.3, Results, model predictions of behavioural threshold, dynamic range and intensity difference limen in implant users are presented. The predictive accuracy of the psychophysical model and the effects of stochastic activity on those predictions are explored. In Section 6.4 the implications of stochastic activity in AN response for cochlear implant research are discussed.

## 6.2 Methods

### 6.2.1 Psychophysical model

The form of the psychophysical model is given in Figure 6-1. This model follows the approach of McKay and McDermott (1997), O’Leary et al. (1995), White (1987) and White (1984a), in which loudness is related to the response of the AN to an electrical stimulus. Such models can be derived from psychophysical models of acoustical stimulation which relate loudness to AN response (e.g., Oxenham and Moore, 1994; Green and Swets, 1973). Here it is assumed that the peripheral sections of acoustical psychophysical models can be attributed to excitation of the AN and the remaining sections attributed to more central processing. To construct the model, the excitation sections are replaced by a model of electrical stimulation of the AN. The central sections of the model are similar to the central sections used to model normal hearing. Accordingly, the model has two major divisions, an *auditory nerve* section and a *psychophysical* section.

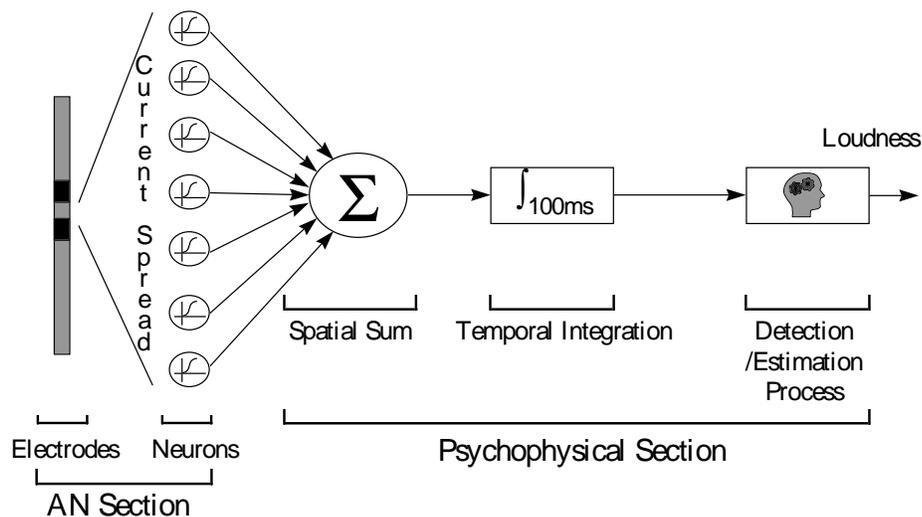


Figure 6-1: Cochlear neural model and behavioural loudness model.

#### AN section

In this chapter, two models of the AN neurons are compared: a stochastic model and a deterministic model. For the AN section of the model a description of total AN response to electrical stimulation based on the model of White et al. (1987) and White (1978) is utilised, which was developed further

in Chapter 5. Parameters for both the stochastic and the deterministic model were determined. The distributions of model neuron parameters are set to approximate those seen in physiological data from cat (see Chapter 5). The output of the total auditory nerve model is a set of probabilities of discharge, one discharge probability for each neuron in the model cochlea. The central component (i.e., the spatiotemporal spike counter) of the model is developed in such a way that for a given stimulus its output is a single mean and a single variance, where the distribution is known. This allows the direct use of signal detection theory to predict the performance of cochlear implant users for a range of psychophysical measures of loudness perception.

The neural excitation model is dependent on a range of different stimulus parameters which have been shown to be important in the perception of loudness. It takes into consideration the stimulus intensity, phase duration (pulse width), pulse rate and pulse train duration. It also incorporates the two most common types of electrode configuration, monopolar and narrow-bipolar. Other configurations could easily be modelled.

### Psychophysical section

Following O’Leary et al. (1995) and White (1984a), loudness is considered to be a monotonic function of the total discharge rate of all auditory nerve fibres activated by the electrical stimulation. Following White (1984a), from psychophysical measurement of threshold versus stimulus duration (Shannon, 1993; White, 1984b), a 100 ms rectangular integration window is used<sup>1</sup>. These two aspects, the summing of responses across all neurons and the long-term temporal integration, can be modelled by spatial summation and temporal integration of the output of the neural section.

**Remark 6.2.1** If the model were to be extended to investigate psychophysical results for higher pulse rate stimuli, short-term integration effects could also be included in the psychophysical section using the approach of McKay and McDermott (1997). Note also that for higher pulse rates, in order to take into consideration refractory effects, the pulse-train model derived in Chapter 7 should be used in the AN section of the model in place of the single-pulse model of Chapter 5 used in this chapter. See Chapter 7 for a discussion of further inter-pulse effects which may need to be considered also.

The output of the entire model is therefore a prediction of the mean and variance over a 100 ms period of the total AN response to a particular electrical stimulus. Signal detection theory can

---

<sup>1</sup>If the stimulus is < 100 ms, as is the case for almost all the data presented in this chapter, the integrator’s output is simply the sum of all spikes that occur during the 100 ms integration period. For the cases where the stimulus is > 100 ms (Sections 6.3.6, 6.3.7 and 6.3.8), the output is the sum of the responses elicited by the maximum number of pulses which can appear within the temporal integration window. For a uniform pulse train this is equal to the elicited response to  $n$  pulses, where  $n$  is the pulse rate times 100 ms.

then be used to predict specific detection/estimation tasks, such as behavioural threshold, dynamic range and Intensity Difference Limen (IDL).

### 6.2.2 Determination of psychophysical measures

The output of the neural/psychophysical model developed in Section 6.2.1 can be used to predict psychophysical results by either Monte Carlo or analytical methods. Monte Carlo techniques are useful because they can use the same paradigm with which psychophysical results are collected. They are also more accurate because they use the actual response distribution of the model. Analytical/numerical methods on the other hand are more efficient computationally, but require approximation of the response distribution.

The following describes the various methods used in determining the model's performance in a number of different psychophysical tasks.

From Section 6.2.1, the loudness of two stimuli are deemed to be equal when the output of the temporal integration section of the model is identical for the two stimuli. The loudness estimated by the model is assumed to be monotonically related to the output of the temporal integration section of the model.

#### **Intensity discrimination: Threshold and IDL**

A range of psychophysical paradigms were used in the collection of the behavioural thresholds and IDLs presented in Section 6.3, many of which were not criterion-specific. Therefore, to simplify the prediction of these data, a two-interval forced-choice (2IFC) paradigm is chosen as a standard procedure to assess the model's prediction of these data. Following the definition of equal-loudness, the decision rule is that the interval creating the greater number of discharges is the one containing the stimulus, in the case of a threshold measurement, and is the interval containing the louder stimulus, in the case of an IDL measurement. An ideal observer is assumed. Because it is a forced-choice paradigm, an unbiased guess is made if the number of discharges is equal for the two stimuli. Both Monte Carlo and analytical methods can be applied to provide measures of threshold and IDL for the model.

In Monte Carlo applications of the model, pseudo-random number generation is used to predict the number of neurons which respond to a given pulse. For each neuron a pseudo-random, uniformly-distributed number between 0 and 1 is generated, and if the number is less than the probability of discharge for that neuron, as determined by the model statistics and the stimulus parameters, the neuron is deemed to have responded. This is repeated for each pulse in the pulse train. The number of responses to each pulse is then passed through the psychophysical section of the model to give the total number of responses occurring within the temporal integration period. This output can then be used directly in any criterion-specific psychophysical paradigm.

In a forced-choice paradigm, this procedure is carried out once for the comparison stimulus and repeated a number of times for the reference stimulus, the number of repetitions depending on the number of intervals desired, e.g., once for a two-interval procedure. For a comparison stimulus of greater intensity than the reference stimulus, if the number of responses to the comparison is greater than the number of responses to each of the reference intervals, then the model is deemed to have correctly identified the comparison. If any one of the reference intervals produces a response greater than the comparison, then the model is deemed to have chosen incorrectly<sup>2</sup>. Using this approach, any criterion-specific paradigm, such the up-down procedures reviewed by Levitt (1970), can be simulated.

In analytical applications of the model, the distribution of the model's output is approximated by an analytical expression. Then signal detection theory is used to derive an analytical expression to predict the model's performance of any criterion-specific psychophysical detection/discrimination task. Expressions for predicting the model's performance of a 2IFC paradigm are derived.

**Lemma 6.2.1** Consider two discrete random variables  $X_1$  and  $X_2$  which describe the number of discharges, as predicted by the model developed in Section 6.2.1, produced in response to stimuli 1 and 2 respectively.  $X_1$  and  $X_2$  have probability mass functions  $f_1$  and  $f_2$  with means of  $\mu_1$  and  $\mu_2$  and standard deviations of  $\sigma_1$  and  $\sigma_2$  respectively.  $f_1$  and  $f_2$  have values for all integers within the bounds  $[0, X_{\max}]$ , where  $X_{\max}$  is the maximum number of discharges possible.

If stimulus 2 has the greater intensity, the probability of the detection/discrimination system choosing correctly that it has a greater intensity than stimulus 1 is then:

$$\begin{aligned} \Pr\{\text{Choosing correctly}\} &= \sum_{n=0}^{X_{\max}} f_1(n) \sum_{m=n+1}^{X_{\max}} f_2(m) \\ &+ \frac{1}{2} \sum_{n=0}^{X_{\max}} f_1(n) f_2(n) \end{aligned} \quad (6.1)$$

**Proof:** see Appendix C.

**Remark 6.2.2** From Chapter 5:

If  $\mu_i < 15$ ,  $f_i$  is well approximated by the Poisson probability mass function (p.m.f.):

$$f_i(x; \mu_i) = \frac{e^{-\mu_i} \mu_i^x}{x!} \quad x = 0, 1, 2, \dots \quad (6.2)$$

---

<sup>2</sup>If the number of responses is equal for the comparison and *all* the reference intervals, and consequently no decision can be made, then another pseudo-random, uniformly-distributed number between 0 and 1 is generated, and if the number is less than the inverse of the number of intervals (0.5, in the case of the two-interval paradigm), then the model is deemed to have guessed the correct interval.

or if  $\mu_i \geq 15$ ,  $f_i$  is well approximated by the Gaussian probability density function (p.d.f.):

$$f_i(x; \mu_i, \sigma_i) = \frac{1}{\sqrt{2\pi}\sigma_i} e^{-(x-\mu_i)^2/2\sigma_i^2} \quad -\infty < x < \infty \quad (6.3)$$

Lemma 6.2.1 assumes both  $X_1$  and  $X_2$  are discrete random variables having integer values bound by  $[0, X_{\max}]$ .

Two possible changes can be made in order to use Equation 6.3 for the case where  $\mu_i \geq 15$ . Either the appropriate sums in Lemma 6.2.1 can be changed to integrals, or the Gaussian p.d.f. can be discretised such that it only has values for integers within the bounds  $[0, X_{\max}]$ , forming a p.m.f. with the same formula as Equation 6.3. These methods produce practically identical results, but the second method has been used in the model predictions, because it is computationally faster.

Sample psychometric functions of behavioural threshold generated using this method are plotted in Figure 6-2. These were obtained from a model of 10 000 neurons (see Section 6.2) in response to a single 100  $\mu\text{s}$ /phase biphasic pulse generated with a monopolar electrode configuration. The deterministic version of this model produces a step-shaped psychometric function, which, because of the lack of any variance in the response, corresponds to one more neuron responding to the comparison stimulus than to the reference stimulus (no stimulus when determining behavioural threshold). In contrast, the stochastic model produces a psychometric function which rises smoothly from 50% (chance performance level) to 100% discrimination, as the neural responses become more discriminable, i.e., as the overlap of the response distributions of the two stimulus conditions decreases.

A point on the psychometric function which is defined as the threshold (or difference limen) can then be chosen to match the paradigm which was used in the collection of the psychophysical data which are to be predicted. However, not all of experimental methods used to collect the data presented in the Results were criterion-specific. In such cases 70.71% is used as the criterion<sup>3</sup>.

In order to test the accuracy of the analytical estimate, Monte Carlo simulations of a standard up-down threshold procedure (Levitt, 1970) were conducted using the psychophysical model. For each threshold measurement the simulation was run until 10 turning points were reached. Threshold was taken to be the mean of the final 8 turning points. Plotted in Figure 6-3 are the means of 4 threshold measurements (o) for four different pulse durations, with error bars indicating  $\pm 1$  standard deviation. Thresholds obtained via the analytical method (lines) accurately estimate the simulation results.

---

<sup>3</sup>70.71% is the point on a psychometric function which is estimated by a two-down, one-up procedure (Levitt, 1970).

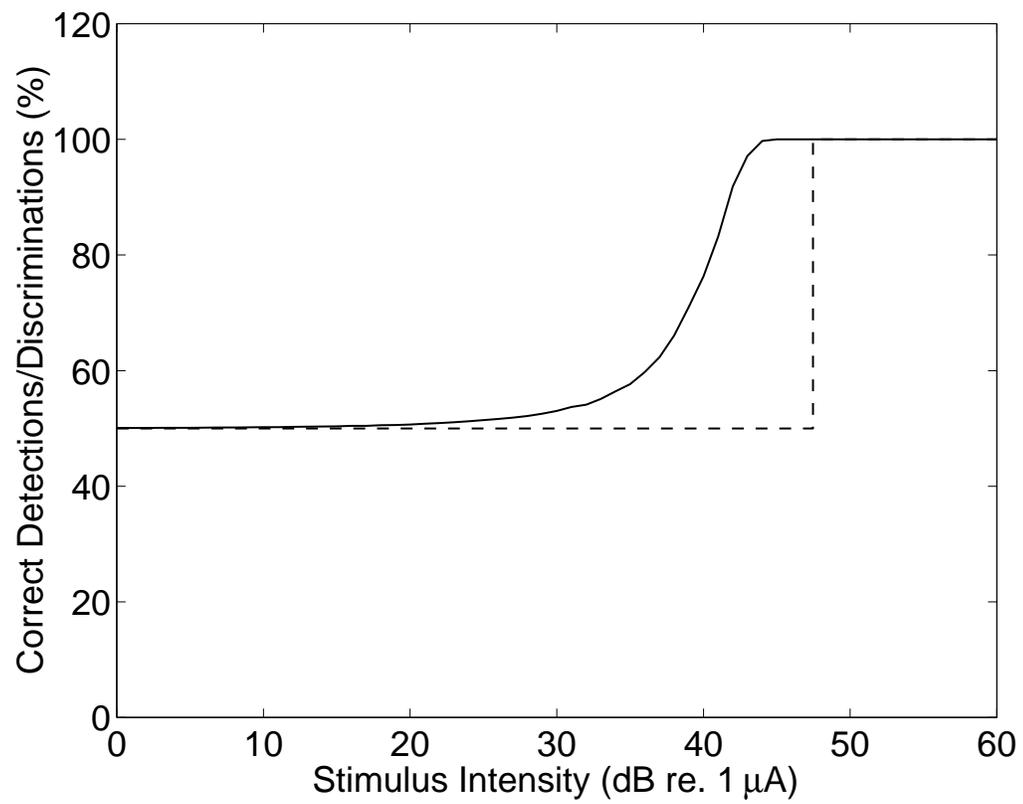


Figure 6-2: Psychometric functions for behavioural threshold predicted by the deterministic (dotted line) and stochastic (solid line) auditory nerve models.

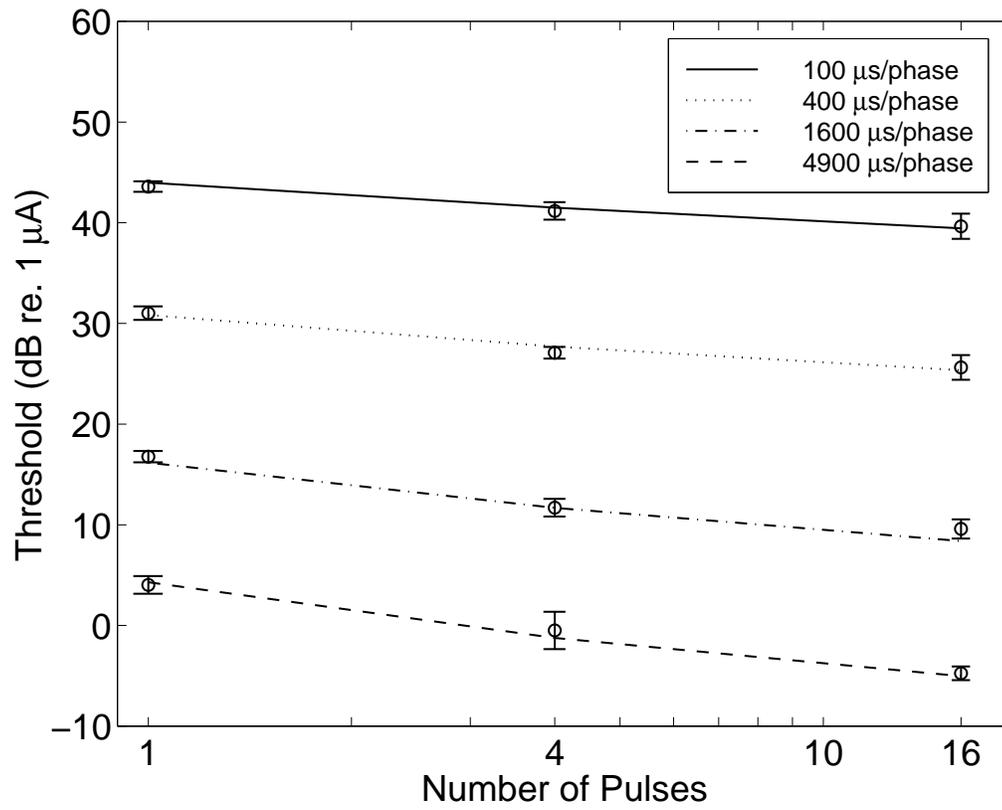


Figure 6-3: Mean thresholds from 4 iterations of Monte Carlo simulations (o), with error bars indicating  $\pm 1$  standard deviation, compared with thresholds obtained via the analytical method (lines).

## Dynamic range

Dynamic range is defined as the intensity difference in dB between behavioural threshold and uncomfortable loudness. Here it is assumed that uncomfortable loudness corresponds to the loudness reaching some fixed level. Because we assume that loudness is monotonically related to the neural response, a fixed loudness corresponds to a fixed number of neural discharges for the whole AN within the period of temporal integration, i.e., the output of the model's temporal integrator section. The number of discharges required to reach uncomfortable loudness will be referred to as  $N_{\text{ucl}}$ . There is no plausible way of determining how many responses correspond to uncomfortable loudness<sup>4</sup>, so predictions for three different values of  $N_{\text{ucl}}$  are therefore presented, which appear to best account for the psychophysical data.

### 6.2.3 Psychophysical data

In Section 6.3.2 threshold versus pulse width data from Pfingst et al. (1991) are presented. In Section 6.3.3 dynamic range versus pulse width data from White (1984b) are presented. In Section 6.3.4 dynamic range versus electrode configuration data from Battmer et al. (1993) are presented, plus previously unpublished data collected by White (Bruce, White, Irlicht, O'Leary and Clark, 1997). In Section 6.3.5 threshold versus electrode configuration data from Smith and Finley (1997) and Miller et al. (1995) are presented. In Section 6.3.6 previously unpublished threshold versus number-of-pulses data collected by White (Bruce, White, Irlicht, O'Leary and Clark, 1997) are presented. In Section 6.3.7 IDL data from Nelson et al. (1996) are presented.

The implant type, experimental setup and methods used by White to collect the previously unpublished data were identical to those used in White (1984b). In summary, subjects were implanted with scala tympani intracochlear electrode arrays of sixteen wires, which were connected to a stimulus generator via a percutaneous link. Bipolar electrode pairs were oriented approximately radially relative to the axis of the cochlea. A monopolar electrode configuration was achieved by using just one electrode from the bipolar pair as the active electrode and an ear-clip located on the ipsilateral earlobe as the return electrode. Thresholds were measured with a modified Békésy tracking procedure using a minimum of 6 threshold crossings for each threshold estimate. The average of the stimulus minima and maxima was computed to determine the estimated threshold stimulus current. Uncomfortable loudness was determined by initiating the stimulus at an amplitude slightly above threshold and gradually increasing the stimulus level on a linear amplitude scale until the subject indicated that the stimulus had reached the maximum loudness desired.

---

<sup>4</sup>Unfortunately, signal detection theory cannot be used to explain the nature or behaviour of uncomfortable loudness. It is possible that other neural mechanisms contribute to the sensation of uncomfortable loudness.

## 6.3 Results

Using the model and methods described in Section 6.2 it is possible to determine threshold, dynamic range and intensity difference limen as a function of stimulus parameters such as phase duration (pulse width), electrode configuration and number of pulses. Before the model's predictions of psychophysical data are presented, it will help to understand the model's neural response as a function of these parameters.

### 6.3.1 Model response properties

Neural response versus stimulus intensity predicted by a model of 10 000 fibres in response to a single pulse of duration 100 and 2 000  $\mu\text{s}$ /phase is plotted in Figure 6-4. The number of discharges is a monotonic function of the stimulus intensity, where there is horizontal shift of the function dependent on the phase duration.

Plotted with a linear ordinate (Figures 6.4(a) and 6.4(c)), the deterministic model and the stochastic model appear to predict similar mean response growth curves. However, when plotted with a log ordinate (Figures 6.4(b) and 6.4(d)), it can be seen that the models predict different neural mean response for pulses that elicit fewer than 1 000 spikes. Not only is there a horizontal shift, but there is also a difference in the slopes of the functions and in the amount of variance. It is likely that auditory threshold is reached with a much smaller number of responses, and hence some or all of a cochlear implant's operating range may be in the region where the stochastic activity is significant. From this it can be expected that the deterministic and the stochastic models will have differing predictions of loudness perception as a function of phase duration and that these differences will be greater at lower stimulus intensities.

Neural response to a single pulse predicted for monopolar (solid line) and bipolar (dashed line) stimulation is plotted in Figure 6-5. For the deterministic model, the recruitment of fibres is steeper for a monopolar stimulus than for a bipolar stimulus, because of the comparatively wider current spread (see Figure 5-12). However, for the stochastic model, this effect is limited to high stimulus intensities and is less pronounced. Again, a significant proportion of a cochlear implant's operating range may lie within the region below 50 dB, where the mean responses of the models diverge and where the variance is significant in the stochastic model. From this it can be expected that the deterministic and the stochastic models will have differing predictions of loudness perception as a function of electrode configuration and that these differences will be greater at lower stimulus intensities.

Furthermore, because of the difference in the response growth curves of the deterministic and stochastic models, (i) predictions of perception as a function of the number of pulses will differ for the two models, and (ii) predictions of intensity difference limen will differ.

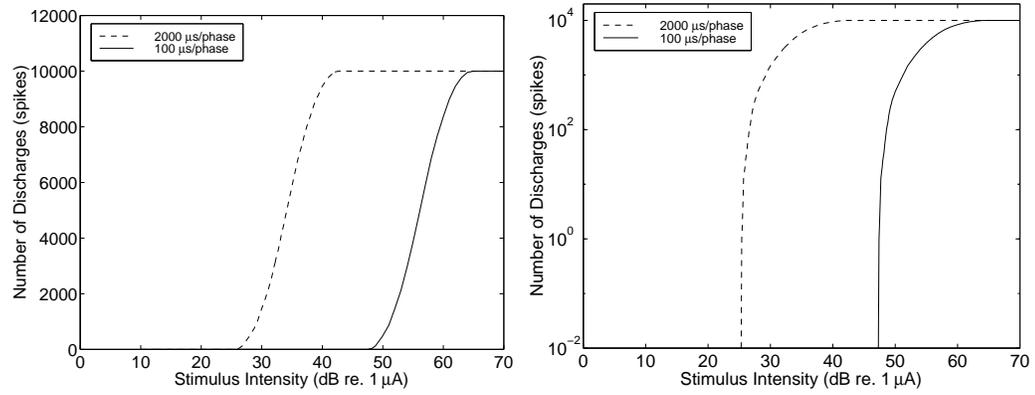
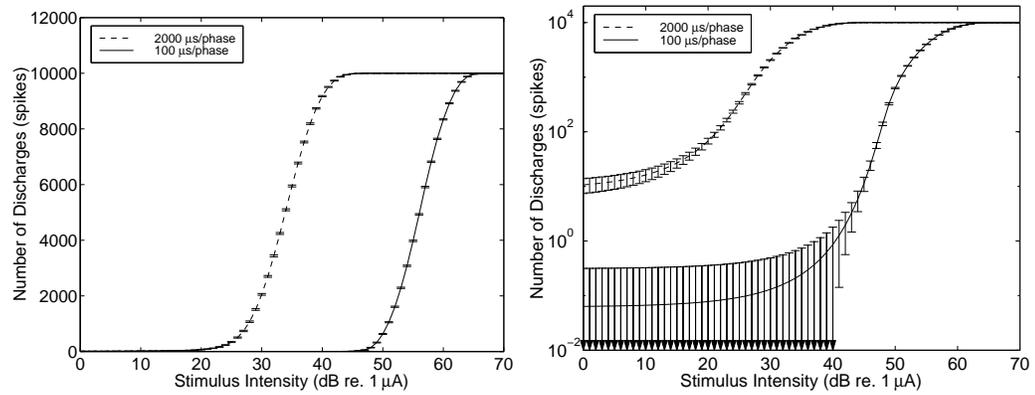
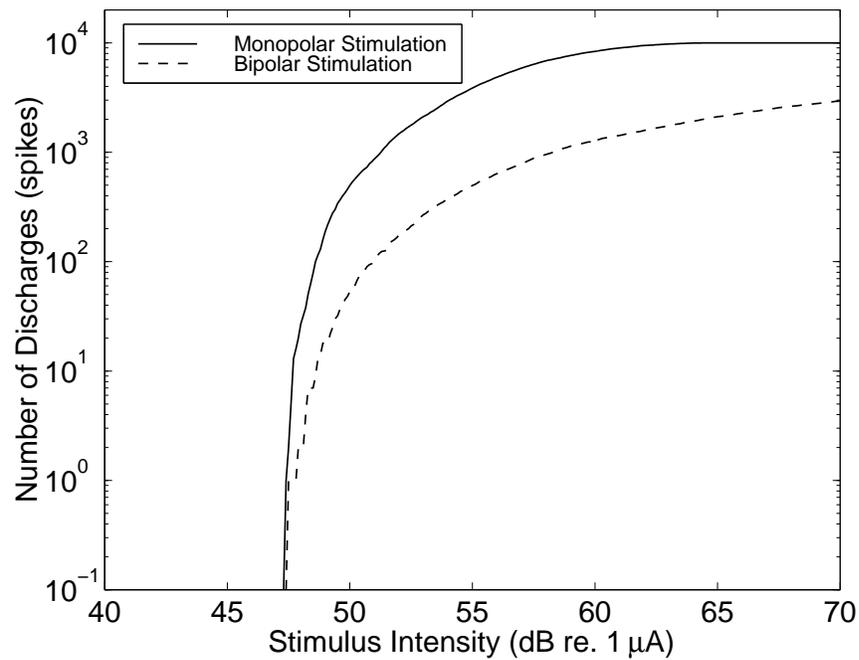
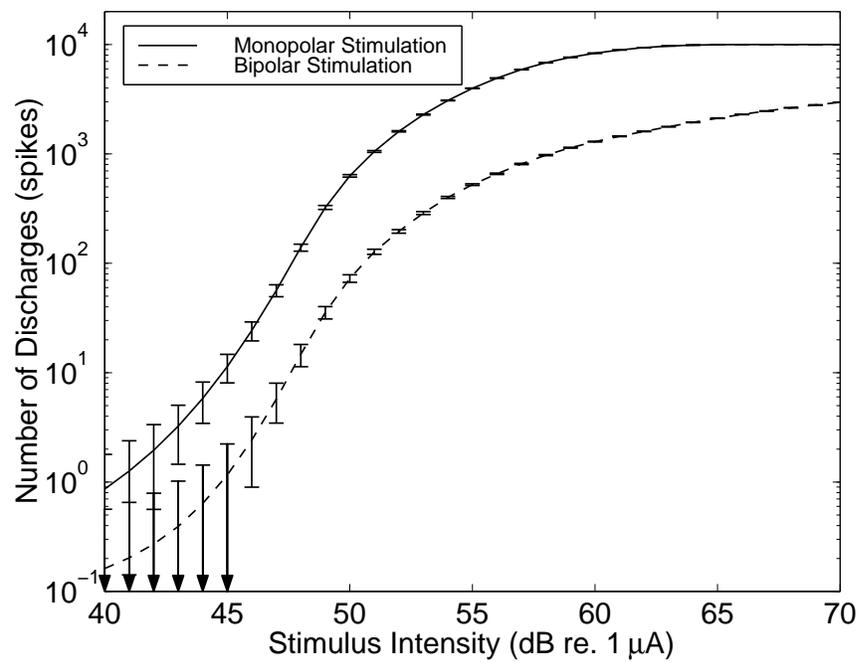
(a) Deterministic model: *linear vertical scale*(b) Deterministic model: *logarithmic vertical scale*(c) Stochastic model: *linear vertical scale*(d) Stochastic model: *logarithmic vertical scale*

Figure 6-4: Neural response versus stimulus intensity from model of 10 000 fibres in response to a single pulse of duration 100 and 2 000  $\mu\text{s}/\text{phase}$  (solid lines and dashed lines respectively).



(a) Deterministic Model



(b) Stochastic Model

Figure 6-5: Neural response versus stimulus intensity from model of 10 000 fibres in response to a single pulse for monopolar (solid line) and bipolar (dashed line) stimulation.

The model's predictions of loudness perception are now compared directly with actual psychophysical data.

### 6.3.2 Threshold versus phase duration (pulse width): Strength-duration

Pfingst et al. (1991) investigated how behavioural thresholds in adolescent or adult male macaques (*M. mulatta* and *M. radiata*) vary as a function of pulse duration for single biphasic pulses. Their data are plotted in Figure 6.6(a) (solid lines), along with a line indicating a 6 dB/octave change in threshold (dotted line) which corresponds to equal charge per pulse. They noted that threshold in some subjects changed by *more* than 6 dB/octave in the region from 1 000 to 2 000  $\mu\text{s}$ /phase, more than would be expected if it was assumed that threshold corresponds to a certain level of charge delivered by an implant. Such steep slopes are commonly observed in human subjects (e.g., see Figure 6.7(a)).

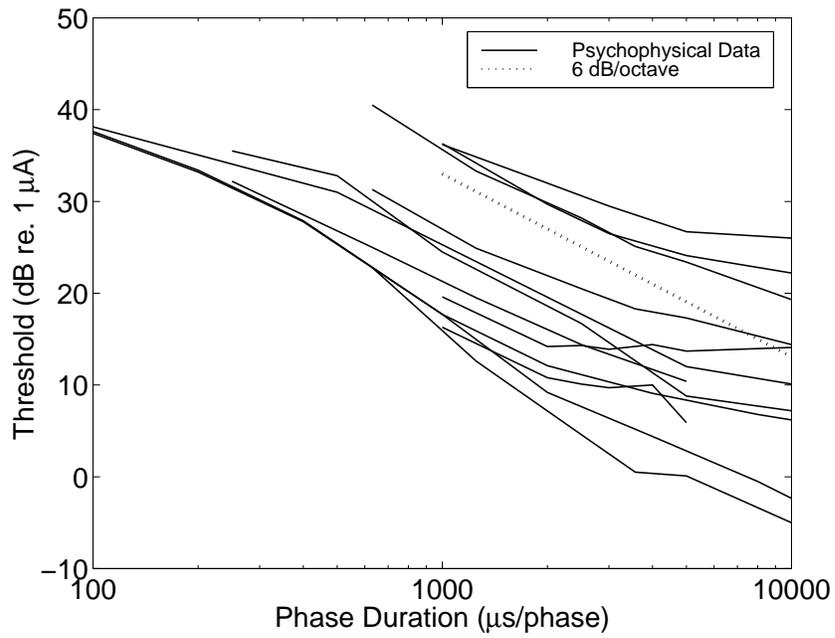
Model prediction of these data are plotted in Figure 6.6(b). Note that the deterministic model predicts slopes of 6 dB/octave at short pulse durations but shallower slopes in the region from 1 000 to 2 000  $\mu\text{s}$ /phase. The stochastic version of the model, in comparison, accurately predicts slopes greater than 6 dB/octave in that region.

### 6.3.3 Threshold and uncomfortable loudness versus phase duration: Dynamic range

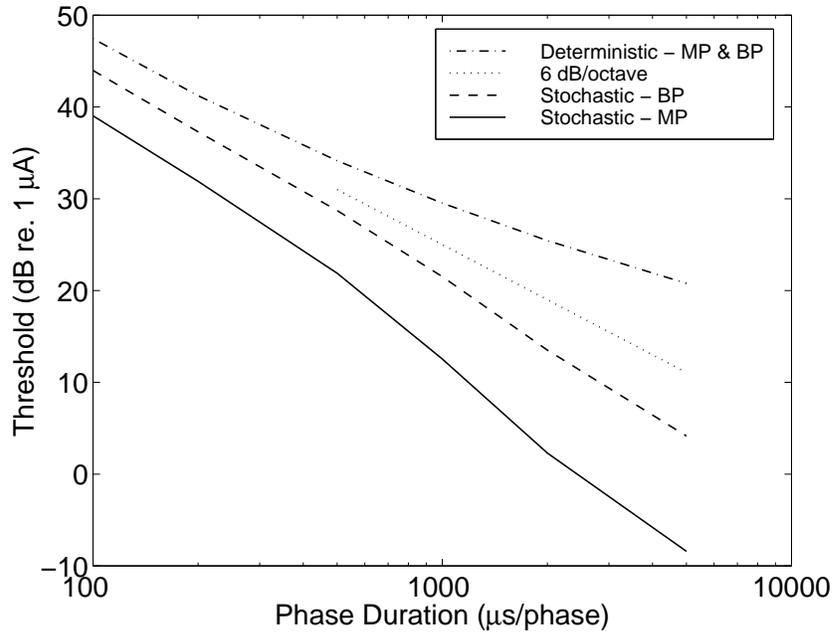
Behavioural threshold and uncomfortable loudness versus phase duration data (White, 1984b) from a human subject are plotted in Figure 6.7(a). The threshold versus phase duration curve has a slope greater than 6 dB/octave in the region from 1 000 to 10 000  $\mu\text{s}$ /phase, which is consistent with the data of Pfingst *et al.* plotted in Figure 6.6(a). In contrast, the uncomfortable loudness curve has a slope of less than 6 dB/octave in this region, causing the corresponding dynamic range, plotted in Figure 6.7(b), to increase with increasing phase duration.

The model predictions of these data are plotted in Figure 6-8. The deterministic model predicts that the threshold and uncomfortable loudness curves will have identical slopes at each phase duration (Figure 6.8(a)), such that the behavioural dynamic range will not change with phase duration (Figure 6.8(b)). In contrast, the stochastic model accurately predicts the changes in slopes of the threshold and uncomfortable loudness curves at each phase duration (Figure 6.8(c)) and the corresponding increase in dynamic range with increasing phase duration (Figure 6.8(d)). It appears that uncomfortable loudness for these data corresponds to an  $N_{\text{ucl}}$  for the model in the region of 100 to 1 000 spikes.

In Figure 6.8(d), the model prediction of uncomfortable loudness for  $N_{\text{ucl}} = 100$  spikes begins to converge with its prediction of threshold. This indicates that threshold for the model corresponds to a higher number of discharges for long pulse durations than for short pulse durations. It can



(a) Psychophysical data



(b) Model predictions

Figure 6-6: Threshold versus phase duration (pulse width) from Figure 2 of Pfungst et al. (1991), and deterministic and stochastic model predictions of these data for monopolar and bipolar stimulation.

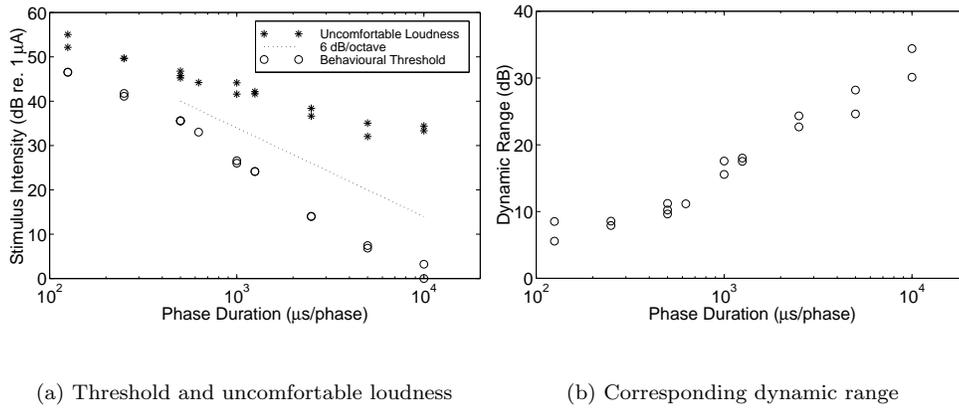


Figure 6-7: Psychophysical data: threshold, uncomfortable loudness and the corresponding dynamic range versus phase duration.

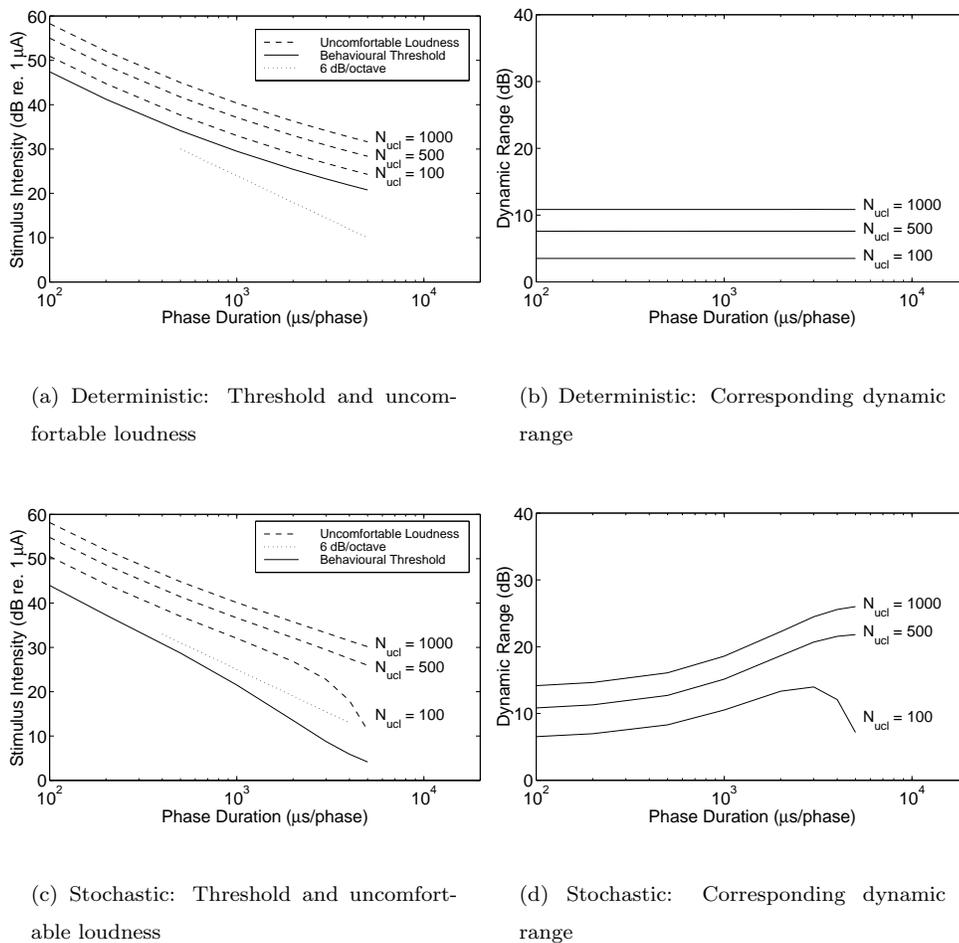


Figure 6-8: Deterministic and stochastic model predictions of threshold and uncomfortable loudness and the corresponding dynamic range versus phase duration.

be postulated that this arises from the inaccuracy of the stochastic single-fibre model at very low stimulus intensities *for long pulses*. Specifically, it can be postulated that discharge probabilities at very low stimulus intensities *for long pulses* are considerably lower than those predicted by the model integrated-Gaussian function (see Chapter 5).

### 6.3.4 Effect of electrode configuration on dynamic range

Behavioural dynamic range data for different electrode configurations are presented here. White (Bruce, White, Irlicht, O’Leary and Clark, 1997) measured behavioural dynamic ranges in response to 1, 2, 4 and 8 pulses of duration 100  $\mu\text{s}$ /phase presented at 50 pps for monopolar and bipolar electrode configurations. Battmer et al. (1993) measured behavioural dynamic ranges in response to pulses of widths 25, 50, 75, 100, 200 and 400  $\mu\text{s}$ /phase for monopolar and bipolar electrode configurations.

These data are plotted in Figures 6-9 and 6-10 respectively, along with model predictions of these data, as BP dynamic range divided by MP dynamic range. All of these data have a ratio of around 1 or less. The deterministic model predicts ratios between 2.7–3.9. In contrast the stochastic model predicts ratios of 0.6–1.2, depending on the value of  $N_{\text{ucl}}$ , which are much closer to the ratios observed in the psychophysical data. Like the dynamic range versus pulse width predictions of Section 6.3.3, it appears that uncomfortable loudness for these data corresponds to an  $N_{\text{ucl}}$  for the model in the region of 100 to 1,000 spikes.

The phenomenon of similar dynamic ranges for monopolar and bipolar electrode configurations has also been observed for stimuli very different from those presented and modelled here. Shannon (1983a) measured behavioural dynamic ranges in response to a 1 000 Hz, 300 ms sinusoidal stimulus for a number of electrodes in a multi-electrode array, using both monopolar and bipolar electrode configurations. Over the entire electrode array the dynamic ranges for monopolar and bipolar stimulation were approximately equal.

### 6.3.5 Effect of electrode configuration on strength-duration

In Figure 6-11, psychophysical strength-duration curves from Miller et al. (1995) and Smith and Finley (1997) for different electrode configurations are plotted. In both data sets the mean thresholds are similar for bipolar and monopolar stimulation at short phase durations, but diverge as phase duration increases. This effect has also been observed in nonhuman primates (Pfungst, Morris and Miller, 1995; Pfingst, Miller, Morris, Zwolan, Spelman and Clopton, 1995).

The change in slope and relative position of the threshold curves with the change in electrode configuration is well predicted by the stochastic model, but not by the deterministic model (see Figure 6.6(b)). Sensitivity testing of the model indicates that the electrode configuration has a larger effect on the predicted threshold curve than either (i) number of fibres or (ii) criterion for the

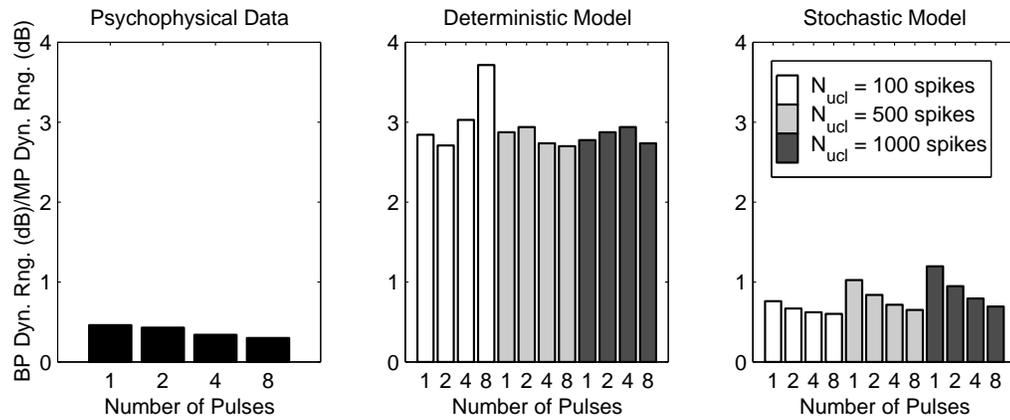


Figure 6-9: BP dynamic range divided by MP dynamic range—psychophysical data from White (Bruce, White, Irlicht, O’Leary and Clark, 1997) and model predictions, for 1, 2, 4 and 8 pulses.

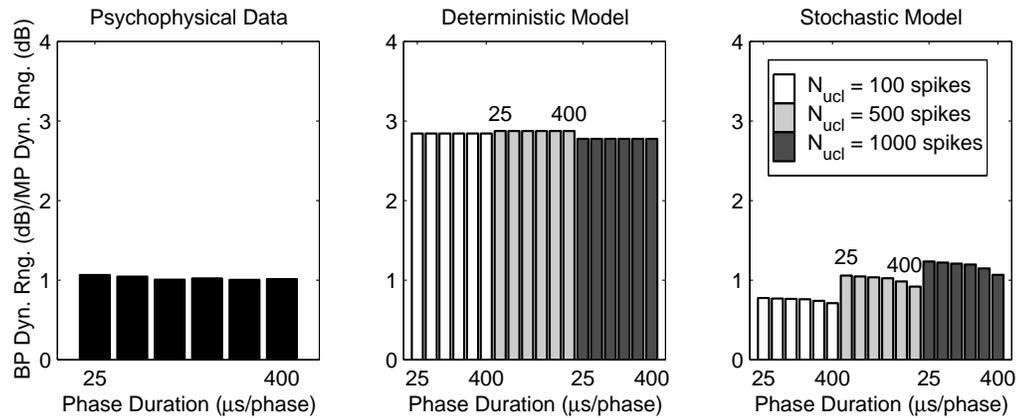


Figure 6-10: BP dynamic range divided by MP dynamic range—psychophysical data from Battmer et al. (1993) and model predictions, for phase durations of 25, 50, 75, 100, 200 and 400  $\mu\text{s}/\text{phase}$ .

psychophysical decision. The range of auditory thresholds in Figure 6.6(a) is also well explained by the range of modes of stimulation used in the collection of those data.

### 6.3.6 Threshold versus number of pulses: Temporal integration

Plotted in Figure 6-12 are thresholds versus number of pulses collected by White (Bruce, White, Irlicht, O’Leary and Clark, 1997) for four different pulse durations, along with deterministic and stochastic model predictions of these data. Unlike the psychophysical data, the slope of the temporal integration curves predicted by the deterministic model (6.12(b)) do not increase with pulse duration, but rather remain at approximately zero. In contrast, the stochastic model (6.12(b)) predicts temporal integration curves with non-zero slopes which increase with phase duration (pulse width).

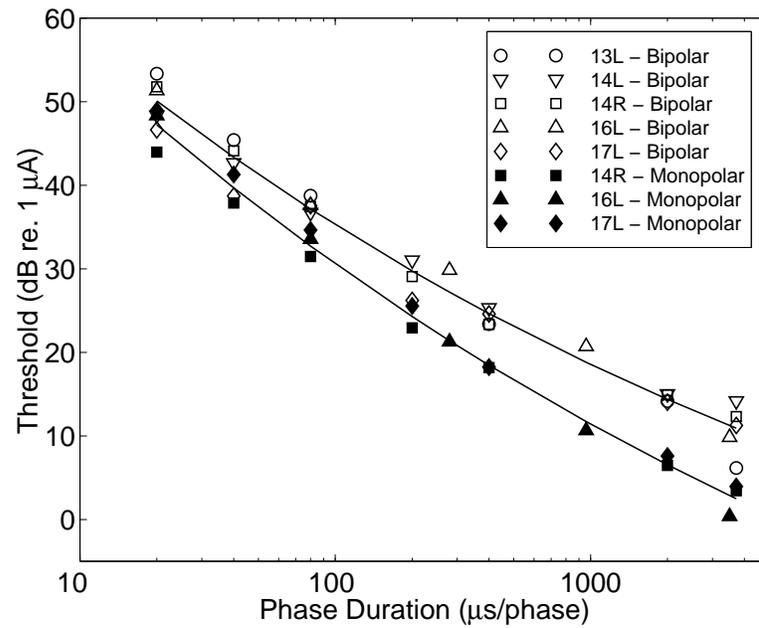
For the longest pulse durations in the psychophysical data (Figure 6.12(a)), there is a sharp increase in the slope of the temporal-integration curves for the larger pulse counts. This is *not* predicted by the model. This may be due to the inaccuracy of the stochastic single-fibre model at very low discharge probabilities for *long pulse durations*. For these very long pulses, model responses for 4 and 16-pulse thresholds are reached at discharge probabilities which it can be hypothesised (see Chapter 5) are not fitted well by the integrated-Gaussian function used in the stochastic model.

### 6.3.7 Intensity Difference Limen (IDL)

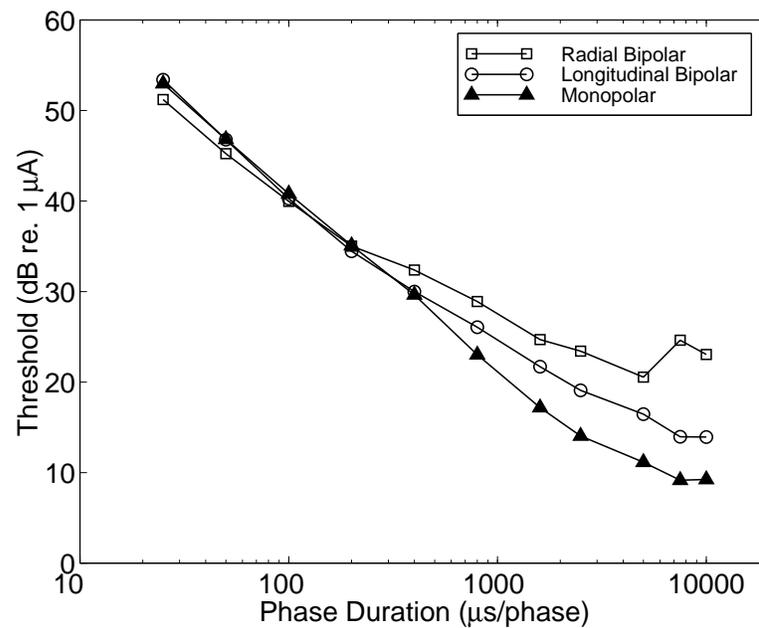
Nelson et al. (1996) measured IDLs in eight cochlear implant users, for 200  $\mu$ s/phase biphasic pulse trains delivered at 125 pps using a bipolar electrode configuration. The psychophysical paradigm used estimated the difference in current level ( $\Delta I_{\mu A}$ ) from the reference current ( $I_{\mu A}$ ) producing a 79.4% correct discrimination. Difference limen were collected and *Weber fractions* expressed in dB ( $= 10 \log_{10}(\Delta I_{\mu A}/I_{\mu A})$ ) calculated for reference current levels spanning each subject’s dynamic range. Linear *Weber functions* ( $= \alpha I_{dB} + b$ ) were fitted to each subject’s data collected over a range of electrode pairs. The reference intensity was expressed as a percentage of the dynamic range ( $\%DR = I_{dB}/DR_{dB} \times 100$ ).

The linear Weber function fits to the psychophysical data of Nelson *et al.* (dotted lines) are plotted in Figure 6.13(a), along with the deterministic model predictions (dot-dashed, dashed and solid lines) of these data. In contrast to the observed behaviour, the deterministic model predicts very erratic Weber functions, with the magnitudes of the predicted Weber fractions all lower than the those seen in the psychophysical data. In agreement with the data, the deterministic model predicts an  $\sim 8$  dB reduction in the Weber fraction over the dynamic range.

The linear Weber function of Figure 6.13(a) are replotted in Figure 6.13(b), along with the stochastic model predictions (dot-dashed, dashed and solid lines) of these data. In comparison to



(a) Guinea pig psychophysical data from Miller et al. (1995). Symbols represent individual data points and solid lines represent least-squares power-function fits to the bipolar and monopolar data.



(b) Cat psychophysical data from Smith and Finley (1997). Symbols and lines represent means across five subjects at each phase duration.

Figure 6-11: Effect of electrode configuration on psychophysical strength-duration curves.

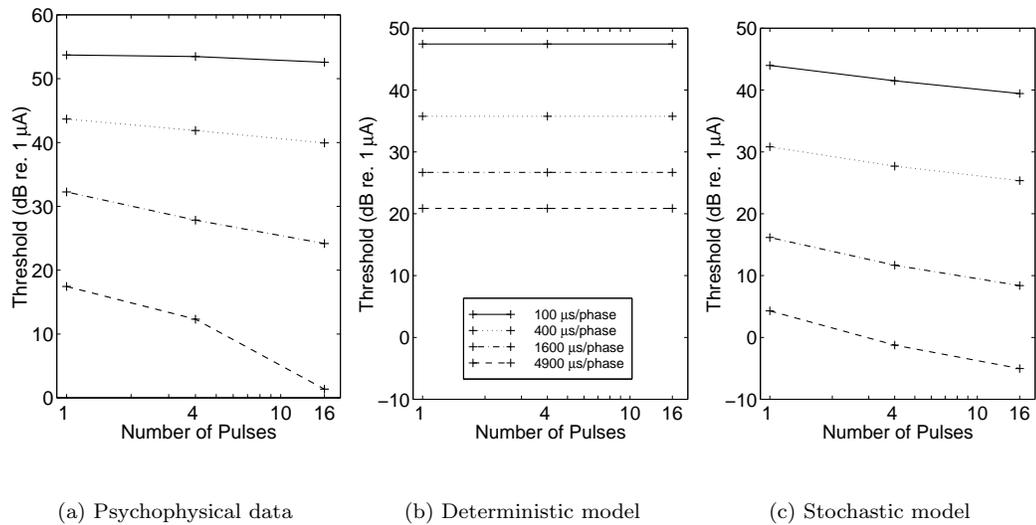


Figure 6-12: Threshold versus number of pulses: psychophysical data and model predictions.

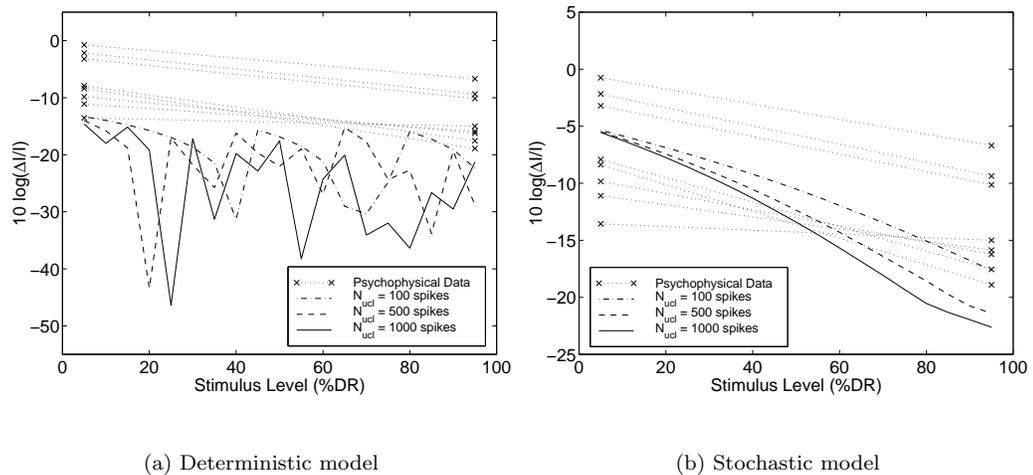


Figure 6-13: Intensity difference limen (Weber fractions) versus stimulus intensity (percentage of dynamic range): linear Weber function fits to psychophysical data from 8 subjects (dotted lines) and model predictions for  $N_{ucl} = 100$  spikes (dot-dashed line), 500 spikes (dashed line) and 1 000 spikes (solid line).

the deterministic model, the magnitudes of the predicted Weber fractions are in better agreement with the psychophysical data. Furthermore, the stochastic model predicts smooth but nonlinear Weber functions, which may provide an even more accurate description of the psychophysical data than a linear Weber function (e.g., see Figures 2 and 6 of Nelson et al. (1996)). Like the dynamic range versus pulse width predictions of Section 6.3.3 and the dynamic range versus electrode configuration predictions of Section 6.3.4, it appears that uncomfortable loudness for these data corresponds to an  $N_{\text{ucl}}$  for the model in the region of 100 to 1 000 spikes.

The stochastic model approximately predicts the slopes of the Weber functions. The reduction in the Weber fraction over the dynamic range is on the order of 10 dB, or greater. Nelson *et al.* found that the average reduction in the Weber fraction over the dynamic range was 8 dB. However, the reduction in the Weber fraction may actually be greater, since shallow Weber functions which contributed to this mean value have been caused by limitations in the intensity resolution of the implanted receiver/stimulators (Nelson et al., 1996).

These predictions can be better understood by observing each model's neural response growth function. In Figure 6-14 neural response predicted by the deterministic model (dashed line) and by the stochastic model (solid line) is plotted as a function of stimulus intensity (percentage of dynamic range, with  $N_{\text{ucl}} = 500$  spikes). It can be seen that:

1. The improvement in intensity discrimination (reduction in Weber fraction) of  $\sim 8$  dB over the behavioural dynamic range is due primarily to the increase in the slopes of the neural response growth curves at uncomfortable loudness when compared with their slopes at threshold. Although the stochastic model has slightly steeper neural response growth at uncomfortable loudness than the deterministic model does, the presence of variance in the loudness indicated by the error bars ( $\pm 1$  std.) causes the discrimination to be worse than if no variance were present, resulting in a difference limen similar to that predicted by the deterministic model.
2. The erratic Weber functions predicted by the deterministic model result from the discrete steps in neural response created by the step-function fits to single fibre I/O functions (see Chapter 5), i.e., the difference limen is purely the increase in intensity required to reach one more fibre's threshold. Therefore, the minimum IDL is practically zero if the reference intensity is immediately below the threshold of the next most sensitive fibre. The maximum IDL corresponds to the widest step in the neural response growth function, i.e., the greatest difference in consecutive thresholds. The smoothness of the Weber functions predicted by the stochastic model is due to the smoothness the stochastic model's neural response growth function, caused by the integrated-Gaussian fits to single fibre I/O functions (see Chapter 5).
3. The increased magnitude of the Weber fractions predicted by the stochastic model, when compared with those predicted by the deterministic model, is due to the variance in the

number of discharges in the stochastic model, indicated by the error bars ( $\pm 1$  std.).

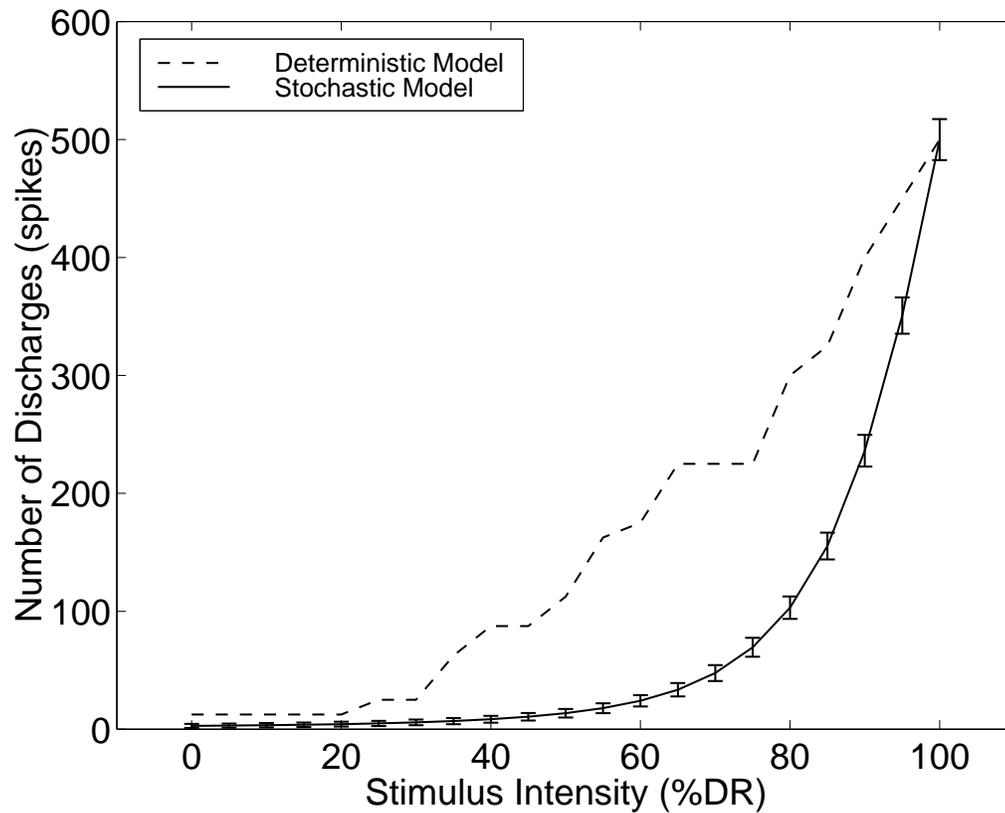


Figure 6-14: Number of neural discharges versus stimulus intensity (percentage of dynamic range, with  $N_{\text{ucl}} = 500$  spikes) predicted by the deterministic model (dashed line) and by the stochastic model (solid line) for a pulse rate of 125 pps. Error bars indicate  $\pm 1$  std.

In Figure 6-14, both the deterministic model and the stochastic model have a lower growth of response near threshold than at higher intensities. If loudness is approximately a linear function of neural response, then the model predicts that loudness growth will be lower at low intensities than at high intensities. This matches the psychophysical data reviewed in Zeng and Shannon (1992).

### 6.3.8 Effect of number of fibres on threshold, dynamic range and intensity difference limen

One of the major forms of inter-subject variability in cochlear implant users is the number of surviving AN fibres. All the model predictions plotted thus far in this chapter were obtained from a model of 10 000 neurons, which is a conservative estimate of the number of surviving auditory nerve fibres in a subject with a profound sensorineural hearing loss (Hinojosa and Marion, 1983). In this section it is investigated how the number of surviving AN fibres affects threshold, dynamic range and IDLs.

Plotted in Figure 6-15 are deterministic and stochastic model predictions of dynamic range (left panes) and Weber fraction (right panes) versus threshold, for bipolar (top panes) and monopolar (bottom panes) stimulation, from models of 1 000 to 30 000 fibres, in response to a 125 pps pulse train. These plots show that both the deterministic and the stochastic model predict changes in dynamic range and Weber fraction with the number of surviving fibres. However, (i) the deterministic model, in contrast to the stochastic model, predicts practically no change in threshold with the number of fibres, and (ii) the deterministic model predicts decreases in dynamic range and Weber fraction with increasing neural survival, whereas the stochastic model predicts the reverse, except for small numbers of surviving fibres in bipolar mode with a high  $N_{uct}$ . For the deterministic model, the intensity required to reach uncomfortable loudness decreases with the number of surviving neurons and threshold remains unchanged, producing a net decrease in the dynamic range. In contrast, both the threshold and the uncomfortable loudness of the stochastic model decrease with increasing neural survival and the rate of decrease of threshold is *greater* than that of uncomfortable loudness, producing a net increase in the dynamic range.

It is not possible to determine the number of surviving fibres in live implant subjects, so it is infeasible to compare these model predictions directly to psychophysical results. However, from Hinojosa and Marion (1983) and similar studies it can be assumed that nerve survival would vary considerably in any reasonably-sized sample of implant subjects. In Fig. 7 of Nelson *et al.* (1996), Nelson *et al.* plotted dynamic range and Weber fraction versus threshold for eight human implant patients, using either a bipolar (adjacent electrodes) or a wide bipolar (one separating electrode) configuration. In Figure 6-16 their linear fits to these data are replotted. If it is assumed that neural survival varies in these eight subjects and there are no other significant differences, then it would be expected the linear fits to the psychophysical data to lie somewhere between those predicted for bipolar stimulation and those for monopolar stimulation. Comparison of the model predictions with the psychophysical data suggests that the stochastic model, in contrast to the deterministic model, can account for the psychophysical data reasonably well. If different distributions of single-fibre I/O functions exist in these eight subjects, as suggested by studies of deafened animals (Shepherd and Javel, 1997), this may further explain the range of psychophysical data<sup>5</sup>.

## 6.4 Discussion

### 6.4.1 Using the model to understand psychophysical behaviour

The relationship between cochlear response and behaviour is defined by the central component of the model. A spike-counting model (i.e., perfect spatiotemporal summation) is used, as used

---

<sup>5</sup>Other factors such as electrode placement and the cochlea's anatomy can cause threshold to vary among subjects.

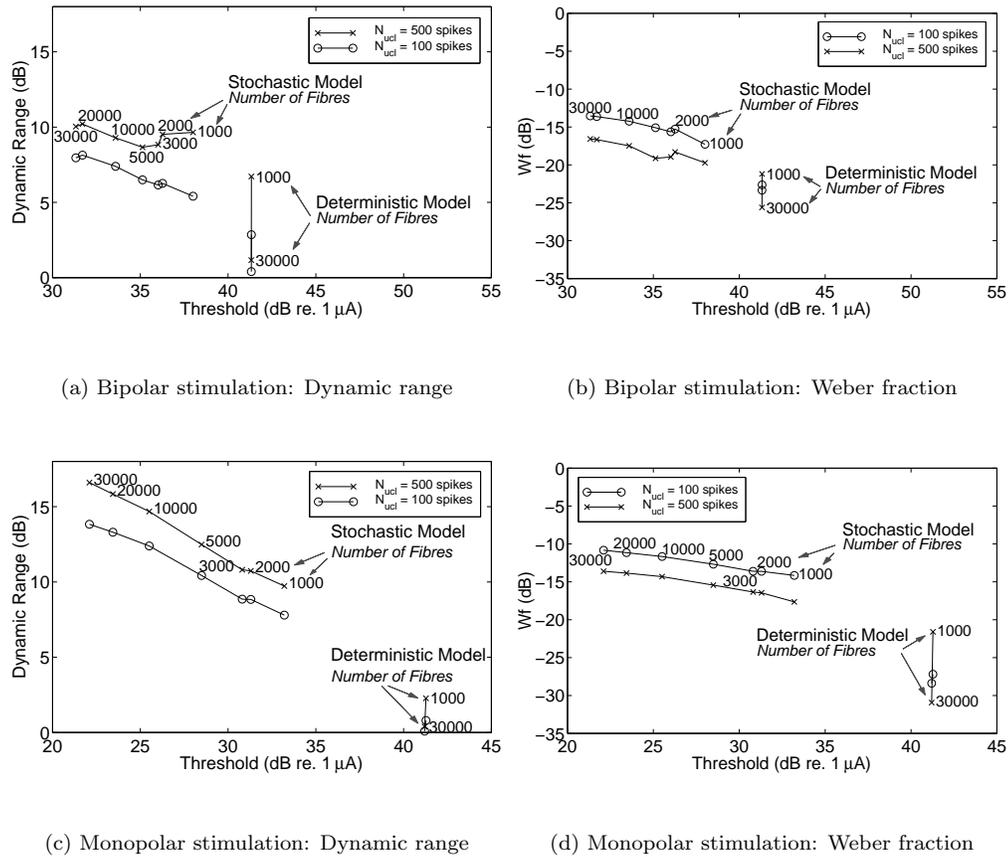


Figure 6-15: Deterministic and stochastic model predictions of dynamic range (left panes) and Weber fraction at 75% of dynamic range (right panes) versus threshold, for bipolar (top panes) and monopolar (bottom panes) stimulation, from models of 1 000 to 30 000 fibres, in response to a 125 pps pulse train. The number of fibres used in the model to generate each point are as labelled. Curves are plotted for two different values of  $N_{ucl}$ , as indicated in the figure legend.

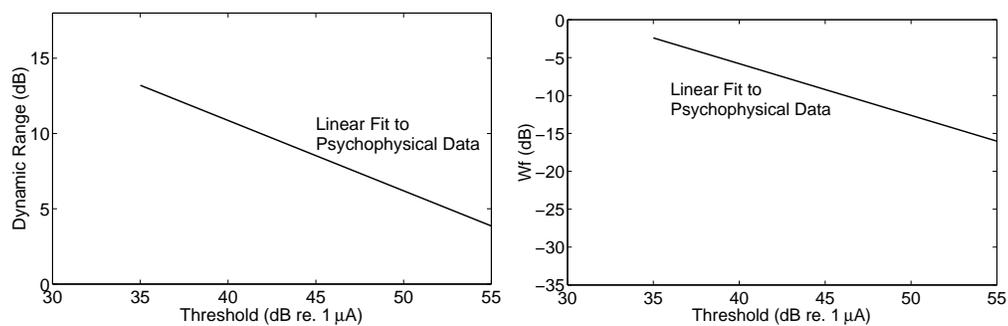


Figure 6-16: Linear fit to psychophysical data from Fig. 7 of Nelson et al. (1996): corresponding dynamic range (left pane) and Weber fraction at 75% of dynamic range (right pane) functions for a 125 pps pulse train.

in some models of normal hearing. This model is relatively simple to use. For example, (i) by adjusting the amplitude of a stimulus until the elicited cochlear output is equal to that elicited by another stimulus, the two stimuli can be made to evoke the same loudness; and (ii) the output is a single mean and a single variance for each stimulus, so relatively simple statistical methods can be used to measure detectability or discriminability of stimuli.

Understanding how the cochlear response is affected by stimulus parameters, electrode configuration, nerve pathology, and nerve survival is necessary for understanding how these variables control behavioural response. The cochlear response is controlled by 3 primary factors: (i) The individual fibre I/O functions—it has been shown that model predictions for a step-function (deterministic model) and for an integrated-Gaussian function (stochastic model) are very different<sup>6</sup>; (ii) Differences in fibres—the distribution of fibre I/O slopes is broad and unimodal. The deterministic model cannot incorporate the slopes of I/O functions, let alone their distribution. In contrast, the stochastic model is able to describe the distribution of slopes, such that fibres with shallow slopes contribute the majority of cochlear activity at low stimulus intensities. In sharp contrast, fibres with steep slopes contribute the majority of cochlear activity at high stimulus intensities. In addition, thresholds also vary (independently of the slope) among fibres. (iii) The large impact of phase duration (pulse width) on the fibre's discharge probability function.

The impact of the 3 primary factors, plus other factors such as the number of pulses, electrode configuration and nerve survival, can be understood and predicted. For example, in order to maintain the same level of audibility when the number of stimulus pulses is halved, pulse amplitude must be increased so that the average discharge probability is doubled. In a similar manner, in order to maintain the same loudness when switching from monopolar to bipolar stimulation, pulse amplitude must be increased to exactly compensate for the reduction in discharge probability across the cochlea. Similarly, a cochlea with poor nerve survival requires a compensatory increase in pulse amplitude. Furthermore, the relative noise level of fibres may be quite different for different cochlear pathologies (Shepherd and Javel, 1997). This directly impacts the slope of individual fibres and therefore the slope of total cochlear response. For example, in subjects with a pathology that causes fibres to be relatively noisy, one would expect dynamic ranges to be larger and intensity discrimination to be poorer, particularly at low loudness.

## 6.4.2 Applicability of the model

The results presented in this chapter have important consequences for physiological studies, for investigation of neural sound coding and for speech processing strategies. For physiological studies,

---

<sup>6</sup>Because AN fibres may only discharge in response to only a fraction of the pulses in the stimulus, discharge probabilities for individual fibres are best visualised using a 2-3 decade logarithmic scale. Cochlear response functions (e.g., Figure 6-4) should be viewed over a much larger range (e.g., 6 decades)!

the results suggest that a simple deterministic threshold measurement is not a sufficient description of neural response to electrical stimulation, but rather discharge probability functions are required. Such measurements not only provide a more accurate description, but also emphasise how stimulus intensity significantly affects what spatiotemporal patterns of response are occurring in the AN and how information could be coded by these patterns. For psychophysical studies, the results indicate that a number of perceptual measures are better predicted by the stochastic model. The source of this improvement is in the description of the most peripheral section of the auditory pathways. It is therefore suggested that consideration of stochastic activity in AN response may produce better predictions and explanations of a range of other psychophysical measures not investigated in this chapter.

An independent group (Ferguson et al., 1998) has implemented the model developed in this chapter and compared the predictions of threshold as a function of pulse duration for monopolar and bipolar stimulation modes with experimental data. Analysis of data from three species indicated that variance of behavioural thresholds is an increasing function of pulse duration, and that these results are corroborated by the predictions of the stochastic version of the model. The results are not predicted if the stochastic component is not included. This indicates that the importance of stochastic activity in the AN extends beyond the psychophysical data investigated in this chapter.

### 6.4.3 Model extensions

In this chapter a model of loudness in cochlear implants users based on physiological data has been derived and been used to investigate a number of different psychophysical phenomena. In all the cases examined so far, the model predicts the perceptual performance of cochlear implant users significantly better when stochastic activity is included in the neural section of the model.

However, although most of the psychophysical data investigated in this chapter are well predicted by the stochastic model, extensions or revisions of this model may significantly improve predictions and the understanding of the functional significance of the physiology.

As was raised in Chapter 5, the neural section of the model is derived from physiological data collected in *cats*. Despite this, the model closely predicts psychophysical data from monkeys, guinea pigs, cats and human cochlear implant users. Further physiological data will be collected from humans using cochlear implant telemetry and non-invasive electrophysiology which should prove useful in refining the simple model of current spread and neural response. A model of current spread in the human cochlea constructed from human cochlear sections (Cohen et al., 1997) may also help to this end.

Another extension to the model would be to allow for other sources of noise. For instance, the survival of inner hair cells in some subjects could result in some remaining spontaneous activity in the AN. This would affect the amount of noise present in the total AN response. Other sources of

noise may also be present in more central sections of the auditory pathways. The effects of both of these potential noise sources can be included in the psychophysical model if their behaviour is known. Initial investigations into such effects have been conducted, where it was assumed that the additional noise is unaffected by stimulus parameters, which could occur if the noise resulted from spontaneous activity in fibres not excited by the electrical stimulus or from some constant source of central neural noise. The consequence of such noise is that thresholds increase with increasing noise levels. If noise levels are high, the stochastic model enters the region where its output approaches the deterministic model predictions. This may account for those strength-duration curves in Figure 6.6(a) which are higher and flatter. Central noise may account for the relatively high Weber fractions produced by loud stimuli.

Other model deviations which effectively reduce the amount of stochastic activity may also account for such data which lies somewhere between the deterministic model and the stochastic model predictions. For instance, particularly focused current fields or extremely low neural density may cause higher probabilities of firing at stimulus intensities within the behavioural operating range. Because neural responses at high discharge probabilities exhibit relatively little variability, stochastic and deterministic model predictions are similar under such conditions.

The physiological data on which the model is based are from *normal hearing* animals. This model does not take into account the effects which prolonged deafness and aetiology have on the response of AN fibres to electrical stimulation (Shepherd and Javel, 1997). An extension to this study could be to model the effects of various aetiologies on current spread and single-fibre I/O functions and the resulting effects on predictions of psychophysical performance. It must be stressed here that these proposed extensions should be applied to the *stochastic model*, since the amount of noise in the stochastic model can always be reduced or increased appropriately. The significance of the single-fibre input/output functions was also investigated by Ferguson et al. (1998). In their investigation, based on the model derived in this chapter, they also utilised raw input/output data gathered from cats tested in their lab who had also performed the behavioural threshold measures. The results of this analysis indicated that the model predictions are more closely matched to the behavioural data when the raw input/output functions are used than when the integrated-Gaussian fits are used.

In the model developed in this chapter, no nonlinearities are present in the central processing, i.e., the psychophysical section of the model. Introduction of nonlinear models of central processing (e.g., Gregson, 1988) may produce better predictions of some psychophysical data. Furthermore, the approach taken in this chapter has been to avoid “curve fitting”, i.e., finding model parameters which produce the best predictions. This has been done in order to demonstrate that parameter values inferred directly from physiological data may be sufficient to produce reasonable qualitative predictions of psychophysical data from many cochlear implant users. However,

because of the reasonable qualitative accuracy of the model predictions, the next step in these comparisons would be to conduct more quantitative evaluations. These could involve fitting the models to psychophysical data of individual cochlear implant users, in order to investigate how model parameters and the corresponding physiology may explain inter-subject variability.

Only responses to stimulation from a single electrode have been investigated in this chapter. In order to model responses to stimulation from multiple electrodes, refractory effects should be incorporated (using the model derived in the next chapter) when the electrodes are stimulating overlapping populations of fibres, and other loudness summation effects may need to be considered when the neural populations excited do not overlap (McKay et al., 1995; Tong and Clark, 1986; Shannon, 1983b). Furthermore, in this chapter the investigation has been limited to low pulse-rate stimuli. With the pulse-train model developed in the next chapter, an appropriate tool is available for extending this investigation to the prediction of psychophysical data for moderate stimulation rates. However, to develop the model for high pulse rate ( $> 1\,000$  pps) stimulation, neurophysiological data must be collected for a range of discharge probabilities at such high pulse rates. It is possible that stochastic activity in the AN will play an even more important role in predicting and understanding behavioural responses to high pulse-rate stimuli.

## Chapter 7

# A stochastic model of the electrically stimulated auditory nerve: Pulse-train response

### 7.1 Introduction

In Chapter 5, a stochastic model of AN response to single electrical pulses was developed, following the conceptual approach used in White et al. (1987), White (1978), and Verveen and Derksen (1968). In contrast to the poor performance of the deterministic model, the stochastic model accurately predicted discharge probabilities measured in response to single biphasic pulses. In Chapter 6 this model was applied in a psychophysical model of loudness in cochlear implant users. The results showed that the stochastic single-pulse model better predicts a range of psychophysical measures of loudness than the equivalent deterministic model does. In this chapter the model is extended to describe responses to pulse-train stimuli.

The final model, although more complex than the single-pulse model, is still computationally efficient and can be fitted easily to the statistics of AN parameters collected from physiological studies. It is therefore suitable for the calculation of large-scale population responses which could be required for the investigation of sound coding in ensembles of nerve fibres, for the explanation or prediction of psychophysical results, or for the development of speech processing strategies for cochlear implants. Furthermore, this simple model is particularly valuable because it is composed of several ‘conceptual components’ (e.g., a stochastic component, a refractory component). The impact of various modifications or exclusions of the components can be easily explored with such models. In a later section the utility of this feature will be illustrated.

The Methods section provides a description of the analysis of the physiological data and the

methods used to derive the neural model. In Section 7.3 the model developed in Chapter 5 is extended to describe responses to pulse-train stimuli. An analytical approximation of the model is derived for trains of evenly-timed, uniform electrical pulses, and response properties of the model at a range of pulse rates are investigated. In particular, the behaviour of the stochastic model is compared to the same model without the stochastic component. In Section 7.4, the effects of pulse-train stimulation on the model of total AN response developed in Chapter 5 are discussed. Finally, in the Discussion the implications of the stochastic behaviour of AN response for cochlear implant research are discussed.

## 7.2 Methods

### 7.2.1 Physiological data

The physiological data presented in this chapter are single-fibre cat AN data from Javel et al. (1987). Along with that paper's analysis of mean discharge rate in response to trains of evenly-timed, uniform, biphasic current pulses, a previously unpublished analysis of this data set is presented. The variance in discharge rate is estimated by calculating the variance in the number of discharges recorded in response to repeated presentation of a 100 ms pulse train. This analysis is examined further in O'Leary et al. (1998).

### 7.2.2 Model of AN response

The various single-fibre models are defined in the form of electrical circuits. Monte Carlo simulations of these circuits can be carried out on a computer and the response of the model recorded directly. This method requires the least number of simplifying assumptions and approximations, producing the most accurate description of the model's response properties, but is computationally laborious. This is because the stochastic version of the model requires many simulation iterations to accurately determine the distribution of responses.

In order to overcome this problem, appropriate simplifying assumptions and approximations have been determined to enable the development of analytical approximations of the electrical circuit descriptions, without significantly changing the response characteristics. Predictions of single-fibre response can be computed simply and efficiently using these analytical approximations and multiple iterations are not required.

## 7.3 Pulse-train response

In Chapter 5 a stochastic model of single-fibre response to single biphasic electrical pulses was developed. Although investigation of single-pulse responses is useful in the development and ver-

ification of a neural model (see Chapter 5), single-pulse stimuli are not the usual output from cochlear implants. Most cochlear implant speech processing strategies typically utilise trains of pulses at rates higher than 100 pulses per second (pps) and train durations up to hundreds of milliseconds. Therefore, a practical model must accurately describe the neural response to trains of current pulses at a range of pulse rates and train durations.

### 7.3.1 Stochastic model of pulse-train response

In the case of pulse trains, the time-dependent nature of neural response to current pulses means that the response to any one pulse in the train will be dependent on how the neuron has responded to the previous pulses in the pulse train. Such dependencies are commonly referred to as *inter-pulse interactions*.

The major form of inter-pulse interaction seen in physiological data is known as the refractory effect, whereby after an action potential is generated there is some time over which the neuron is either unable or less able to be driven to discharge again (Dynes, 1996; Parkins, 1989; Stypulkowski and van den Honert, 1984; White, 1978). The time over which the neuron is subject to refractory effects is notionally divided into the *absolute refractory period* and the *relative refractory period*. The former occurs immediately after action potential generation, and during this time it is not possible to produce another discharge. Following the absolute refractory period is the relative refractory period, during which the neuron is harder to drive to discharge than when it is outside the refractory period. After the relative refractory period the neuron has returned to its ‘normal’ state and its response is no longer dependent on previous activity.

It is possible to approximate this behaviour by introducing a phenomenological mechanism to the model which describes the end result of the refractory effects on the probability of discharge. A refractory potential  $V_{\text{refr}}$  is added to the single-pulse model of Chapter 5, as shown in Figure 7-1. Following an action potential, the ‘threshold’ with which the stimulus potential is compared is raised over the refractory period by some chosen function.

The refractory function  $V_{\text{refr}}$  will have a value of infinity for the absolute refractory period, and will then generally fall smoothly to zero over the relative refractory period. Therefore two assumptions are made regarding the behaviour of the refractory function.

**Assumption 7.3.1**  $V_{\text{refr}}$  is assumed to be monotonically decreasing, i.e., it falls from its maximum value at the beginning of the refractory period to its minimum value at the end of the refractory period without any transitory increases.

**Assumption 7.3.2** The refractory effects are assumed to be of finite duration, i.e.,  $V_{\text{refr}}$  returns to a value of zero within a finite time.

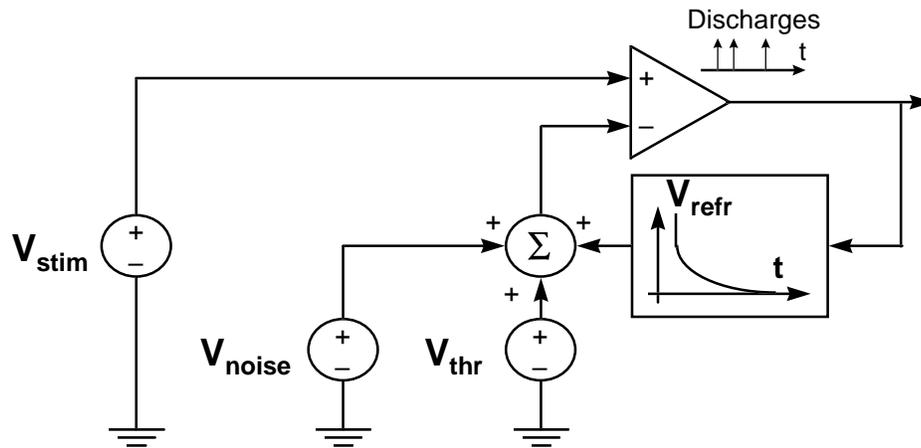


Figure 7-1: Stochastic model of pulse-train response.

Refractory effects may also change the behaviour of the membrane noise. It appears that the magnitude of  $V_{\text{noise}}$  may increase with small depolarisations (see Figure 5.1(b)). The behaviour of the membrane noise during the larger depolarisation which occurs as a result of an action potential is not well understood. This requires a more thorough investigation of the behaviour of ionic channels during and after action potential generation. Some physiological data and modelling results indicate that the noise only increases significantly during depolarisations which do not generate an action potential (see Figure 5-7 of Dynes, 1996), whereas others suggest that the membrane noise could increase greatly during a segment of the refractory period under some conditions (Rubinstein, Matsuoka, Abbas and Miller, 1997; Miller et al., 1997). Due to the uncertainty about the exact nature of these phenomena, two cases will be investigated which are referred to as *Case 1*, where these effects are ignored, and *Case 2*, where the increase in noise during the refractory period is approximated, but only to the extent where the relative noise remains constant during the refractory period.

*Case 1:*  $V_{\text{noise}}$  is proportional to the resting threshold, i.e.,  $\sigma = \text{RS} \times V_{\text{thr}}$ .

*Case 2:*  $V_{\text{noise}}$  is proportional to the refractory-modified threshold, i.e.,  $\sigma = \text{RS} \times (V_{\text{thr}} + V_{\text{refr}})$ .

where Relative Spread (RS) is defined as the standard deviation of the noise ( $\sigma$ ) divided by the threshold. Refer to Chapter 5 for further explanation of the nomenclature.

An excitatory inter-pulse effect has also been observed if the depolarisation caused by a pulse is not great enough to produce an action potential. In this case, any pulse following this subthreshold ‘conditioner’ pulse will have a reduced threshold (Dynes, 1996; Butikofer and Lawrence, 1979). These excitatory effects generally last for approximately half a millisecond, but may have longer

time constants in small diameter fibres (e.g., Frijns, 1995, Chapter IV). This effect requires further investigation, and could be added to the model when sufficient information becomes available.

Auto-correlation of the membrane noise could also produce inter-pulse interactions, i.e., the noise potential during one pulse may be correlated with the noise potential during the preceding and following pulses in the pulse train, such that their responses are not independent. Such effects could be investigated electrophysiologically or by using a Hodgkin–Huxley type model including stochastic activity (Rubinstein, 1995; Hochmair-Desoyer et al., 1984; Lecar and Nossal, 1971), but such an investigation would be extremely difficult either experimentally or computationally and is beyond the scope of this chapter. Therefore a simplifying approximation is used.

**Approximation 7.3.1**  $V_{\text{noise}}$  is considered to be uncorrelated between pulses.

Finally, for the purpose of numerical evaluation the model is discretised, i.e., each pulse  $n$  is divided into a number of discrete bins of equal time. The level of discretisation, i.e., number of bins per pulse, will have an effect at higher pulse rates where a number of pulses fall within the relative refractory period. The number of bins required is dependent on the slope of the refractory function. 10 bins per pulse are used for discretisation level in computing the simulation results, because it was found that increasing the number of bins beyond 10 had negligible effect.

### 7.3.2 Analytical approximation of stochastic model

In this section expressions are developed for the mean and variance of the discharge rate in response to a train of evenly-timed, uniform pulses. In the course of developing these descriptions, an expression for calculating the Interspike Interval (ISI) histogram is also derived. Rather than developing descriptions for the deterministic and stochastic models separately as was done in Chapter 5, it is simpler to derive an expression for the stochastic model and to treat the deterministic model as a special case of the stochastic model, i.e., the case where the noise variance equals zero.

The stochastic model developed in the previous section can be described as a renewal process (Feller, 1968; Cox, 1962). In a renewal process, the waiting times between successive occurrences of an event are mutually independent random variables having the same distribution (Feller, 1968, p. 303). The model qualifies as a renewal process, because after each discharge it begins anew without any ‘memory’ of prior discharges. The theory of renewal processes was originally developed to predict failure times for mechanical or electrical components, or conversely how often a component needs to be replaced or *renewed*. Hence the time between events is called the renewal time ‘ $r$ ’. In this case it is used to model the ISI, which describes the time between consecutive discharges.

Given the mean and variance of the renewal time it is possible to calculate the mean and variance of the discharge rate. Exact expressions for all of these can be derived for the stochastic

model in the case of a pulse train of infinite duration. Therefore, the following approximation is made.

**Approximation 7.3.2** The actual finite pulse train of evenly-timed, uniform pulses is approximated by a pulse train of infinite duration.

**Remark 7.3.1** This approximation constitutes the only difference between the stochastic model suitable for Monte Carlo simulation and the analytical description thereof. The sensitivity of the analytical results to this approximation is examined in Section 7.3.3, where the analytical approximation is compared to results of Monte Carlo simulations.

Following Approximations 5.3.1, 5.3.2 and 5.3.3, all values in the pulse-train model remain constant across pulse  $n$  except the refractory potential  $V_{\text{refr}}$  for *Case 1*, and also the noise potential  $V_{\text{noise}}$  for *Case 2*. The values of  $V_{\text{refr}}$  and  $V_{\text{noise}}$  will depend on the time ‘ $t$ ’ in the pulse train that the last discharge occurred. Therefore,  $E[r]$  and  $\text{var}[r]$  in turn will be dependent on the time of the last discharge. The mean and variance of the renewal time can be calculated using Bayes’ Theorem (Devore, 1987), if the mean and variance of the renewal time dependent on the time of the last discharge can first be determined and the probability distribution is calculated for all the times the previous discharge may have occurred.

That is:

$$E[r] = \int_{\forall t} E[r|\text{last discharge occurred at time } t] \times \Pr \{ \text{last discharge occurred at time } t \} dt$$

and

$$\text{var}[r] = \int_{\forall t} \text{var}[r|\text{last discharge occurred at time } t] \times \Pr \{ \text{last discharge occurred at time } t \} dt$$

For a pulse train of infinite duration consisting of evenly-timed, uniform pulses, the process will reach a state where, from that point on, the probability of response to each pulses does not change. This is known as the *equilibrium state* (Cox, 1962). The mean and variance of the renewal time taken over an *infinite* pulse train will be exactly equal to the mean and variance of the renewal time in the equilibrium state.

Therefore, under Approximation 7.3.2, the case is considered where a discharge has occurred during a pulse while the process is in an equilibrium state. If the discharge has occurred in bin  $j$  of

that pulse, then the expected time to the next discharge is labelled as  $E[r|j]$ . The average renewal time across all values of  $j$  is then:

$$E[r] = \sum_j E[r|j] \Pr \{\text{last discharge in bin } j\} \quad (7.1)$$

and the variance is:

$$\text{var}[r] = \sum_j \text{var}[r|j] \Pr \{\text{last discharge in bin } j\} \quad (7.2)$$

First the mean and variance of the renewal time dependent on the time of the last discharge is calculated.

From Feller (1968):

$$E[r|j] = \sum_{k=0}^{\infty} k f(k|j) \quad (7.3)$$

where  $f(n|j)$  is defined as the probability of the *first subsequent* discharge occurring during pulse  $n$ , given that the last discharge occurred  $n$  pulses previously in bin  $j$ . This scheme is illustrated in Figure 7-2.

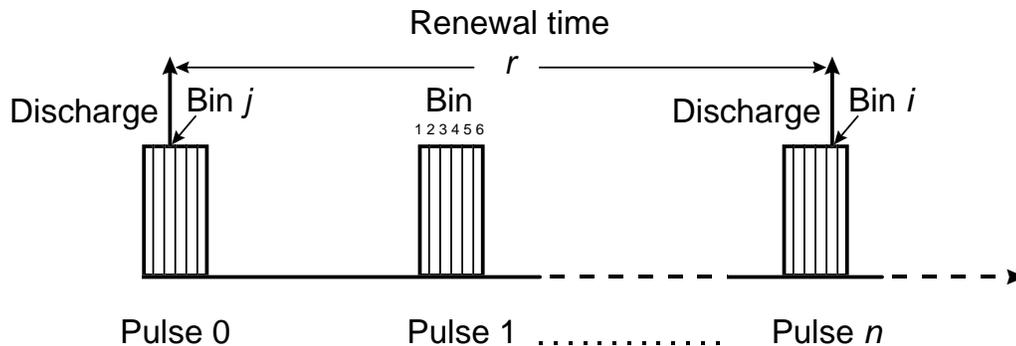


Figure 7-2: Schematic of pulse and bin numbering for calculation of the renewal time  $r$ , where a discharge has occurred in bin  $j$  of pulse 0 and the next subsequent discharge occurs in bin  $i$  of pulse  $n$ . A discretisation level of 6 bins per pulse is shown.

$f(n|j)$  can be calculated by first determining the probability of a discharge occurring during pulse  $n$ , given that the last discharge occurred  $n$  pulses previously in bin  $j$ , which is labelled  $p(n|j)$ .

**Lemma 7.3.1** Given the model of action potential generation described by Figure 7-1; Approximations 5.3.1, 5.3.2, 5.3.3 and 7.3.1; and Assumption 7.3.1: an action potential will occur during pulse  $n$  if  $V_{\text{stim}}(n)$  is greater than or equal to the refractory (and, in *Case 2*, noise) modified threshold by the end of pulse  $n$ .

Define  $g(n_i|j)$  as the probability that  $V_{\text{stim}}(n)$  is greater than or equal to the refractory (and, in *Case 2*, noise) modified threshold in bin  $i$  of pulse  $n$ , given that the last discharge occurred  $n$  pulses previously in bin  $j$ .  $g(n_i|j)$  can be expressed as:

*Case 1:*

$$g(n_i|j) = \frac{1}{2} \left( 1 + \operatorname{erf} \left( \frac{V_{\text{stim}}(n) - V_{\text{thr}}(n) - V_{\text{refr}}(n_i|j)}{\sqrt{2}\sigma} \right) \right) \quad (7.4)$$

*Case 2:*

$$g(n_i|j) = \frac{1}{2} \left( 1 + \operatorname{erf} \left( \frac{V_{\text{stim}}(n)/m(n_i|j) - V_{\text{thr}}(n)}{\sqrt{2}\sigma} \right) \right) \quad (7.5)$$

where

$$m(n_i|j) \triangleq \frac{V_{\text{thr}}(n) + V_{\text{refr}}(n_i|j)}{V_{\text{thr}}(n)}$$

Then, for both cases:

$$p(n|j) = g(n_{\text{final}}|j) \quad (7.6)$$

where  $n_{\text{final}}$  denotes the final bin of pulse  $n$ .

**Proof:** see Appendix D.

Given (7.6), it is possible to calculate the probability of the *first subsequent* discharge occurring during pulse  $n$ , given that the last discharge occurred  $n$  pulses previously in bin  $j$ .

**Lemma 7.3.2** Consider a discrete renewal process, where  $p(n|j)$  is defined as the probability of a discharge occurring during pulse  $n$ , given that the previous discharge occurred  $n$  pulses previously in bin  $j$ .

The probability of the first subsequent discharge occurring during pulse  $n$ , given that the last discharge occurred  $n$  pulses previously in bin  $j$ , is then:

$$f(n|j) = \begin{cases} p(0|j) & \text{for } n = 0 \\ Q(n-1|j)p(n|j) & \text{for } n > 0 \end{cases} \quad (7.7)$$

where  $Q(n|j)$  is defined as the probability of no discharge occurring up to and including pulse  $n$ , given that the last discharge occurred  $n$  pulses ago in bin  $j$ , and is described by the equation:

$$Q(n|j) = \prod_{m=0}^n 1 - p(m|j)$$

**Proof:** see Appendix D.

**Remark 7.3.2** When RS is zero, the model behaves deterministically, and (7.4) and (7.5) become:

$$g(n_i|j) = \begin{cases} 0 & \text{for } V_{\text{stim}}(n) < V_{\text{thr}}(n) + V_{\text{refr}}(n_i|j) \\ 1 & \text{for } V_{\text{stim}}(n) \geq V_{\text{thr}}(n) + V_{\text{refr}}(n_i|j) \end{cases} \quad (7.8)$$

It is possible to reduce the number of values of  $n$  for which it is necessary to evaluate the expressions in Lemmas 7.3.1 and 7.3.2, and thus simplify the calculation of (7.3), because the monotonically decreasing refractory function falls to a value of zero within a finite time (Assumptions 7.3.1 and 7.3.2) and the pulse train is of infinite duration (Approximation 7.3.2).

**Lemma 7.3.3** Following Approximation 7.3.2 and Assumption 7.3.2, consider the case where the pulse train is of infinite duration and there exists some time  $L$  such that refractory effects are negligible for all times  $k \geq L$  such that  $p(k|j) = p, \forall k \geq L$ . Then:

$$\mathbb{E}[r|j] = \sum_{k=0}^{L-1} k f(k|j) + Q(L-1|j) \left( L - 1 + \frac{1}{p} \right) \quad (7.9)$$

and

$$\begin{aligned} \text{var}[r|j] &= \sum_{k=0}^{L-1} k^2 f(k|j) \\ &+ Q(L-1|j) \left( \left( L - 1 + \frac{1}{p} \right)^2 + \frac{1-p}{p^2} \right) - \mathbb{E}[r|j]^2 \end{aligned} \quad (7.10)$$

**Proof:** see Appendix D.

Next the probability distribution is calculated for all the times the previous discharge may have occurred. To do this, a matrix  $\mathbf{B}$  is introduced, where the elements of the matrix,  $B_{ij}$ , are defined as:

$$B_{ij} \triangleq \Pr \left\{ \begin{array}{l} \text{next discharge occurs in any bin } i \\ | \\ \text{last discharge occurred in any bin } j \end{array} \right\}$$

**Lemma 7.3.4** For the renewal process described by Lemmas 7.3.1, 7.3.2 and 7.3.3, the elements of  $\mathbf{B}$  can be calculated via:

$$B_{ij} = \sum_{k=0}^{L-1} Q(k-1|j) p(k_i|j) + Q(L-1|j) \delta_{i=1} \quad (7.11)$$

where  $p(n_i|j)$  is defined as the probability of a discharge in bin  $i$  of pulse  $n$ , given that the previous discharge occurred  $n$  pulses previously in bin  $j$ , and  $\delta_{i=1}$  is one where  $i = 1$  and zero otherwise.

Following Assumption 7.3.1:

$$p(n_i|j) = g(n_i|j) - g(n_{i-1}|j) \quad (7.12)$$

**Proof:** see Appendix D.

Next two column vectors  $\vec{x}$  and  $\vec{y}$  are introduced, where the elements of these vectors,  $x_j$  and  $y_i$  respectively, are defined by:

$$x_j \triangleq \Pr \{ \text{last discharge occurred in any bin } j \}$$

and

$$y_i \triangleq \Pr \{ \text{next discharge occurs in any bin } i \}$$

Following Bayes' Theorem:

$$\vec{y} = \mathbf{B}\vec{x} \quad (7.13)$$

In the equilibrium state the probability that the a spike occurs in bin  $i$  of any pulse is on average equal to the probability that the last spike occurred in bin  $j$  of any pulse when  $i = j$ . Therefore a vector  $\vec{v}$  is defined:

$$\vec{v} \triangleq \vec{y} = \vec{x}$$

Thus (7.13) becomes:

$$\vec{v} = \mathbf{B}\vec{v} \quad (7.14)$$

**Lemma 7.3.5** Given that the probability of discharge is non-zero for all pulses and that the pulse train is of infinite duration (Approximation 7.3.2), the sum of all the elements in each column of  $\mathbf{B}$  equals one, and  $\mathbf{B}$  has at least one eigenvalue of one.

**Proof:** see Appendix D.

Thus,  $\vec{v}$  is the eigenvector of  $\mathbf{B}$  corresponding to the eigenvalue of one. Given that the probability of discharge is non-zero for all pulses and that the pulse train is of infinite duration (Approximation 7.3.2), both  $\sum_j x_j$  and  $\sum_i y_i$  must equal one, and therefore  $\vec{v}$  should be scaled so that  $\sum_j v_j = 1$ .

(7.1) then becomes:

$$E[r] = \sum_j E[r|j] v_j \quad (7.15)$$

and (7.2) likewise becomes:

$$\text{var}[r] = \sum_j \text{var}[r|j] v_j \quad (7.16)$$

All the expressions required to evaluate (7.15) and (7.16) are now available:  $E[r|j]$  can be found from (7.9),  $\text{var}[r|j]$  from (7.10), and the eigenvector  $\vec{v}$  from the values of  $\mathbf{B}$  given by (7.11).

From page 321 of Feller (1968), the mean and variance of the number of discharges  $D$  in a given time window can be calculated from the mean and variance of the renewal time, if the time window is long enough for the distribution of the number of discharges to be well approximated by a normal distribution. For an pulse train of *infinite* duration, the limiting case given by (6.8) of Feller (1968) will be reached. Therefore, given Approximation 7.3.2, the mean ( $E[D]$ ) and the variance ( $\text{var}[D]$ ) of the discharge rate can be described exactly by:

$$E[D] = \frac{1}{E[r]} \quad (7.17)$$

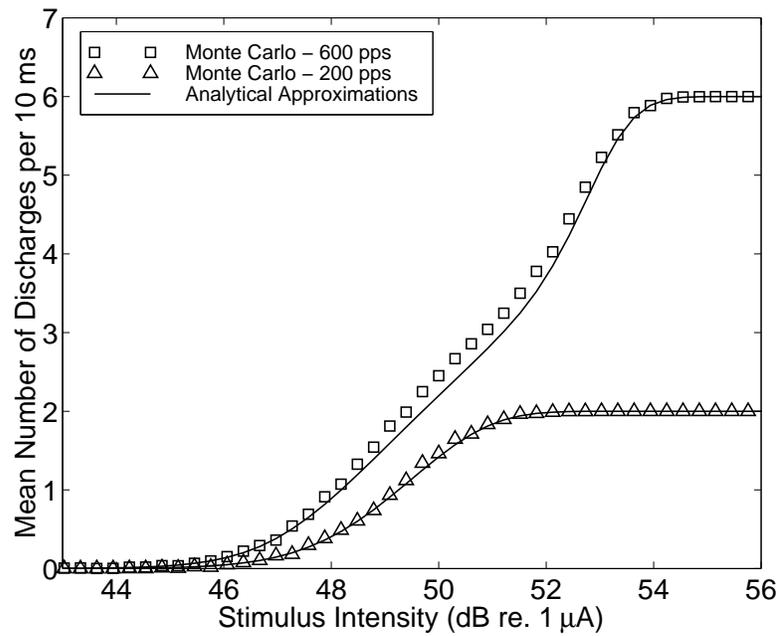
$$\text{var}[D] = \frac{\text{var}[r]}{E[r]^3} \quad (7.18)$$

### 7.3.3 Analysis of the analytical approximation's accuracy

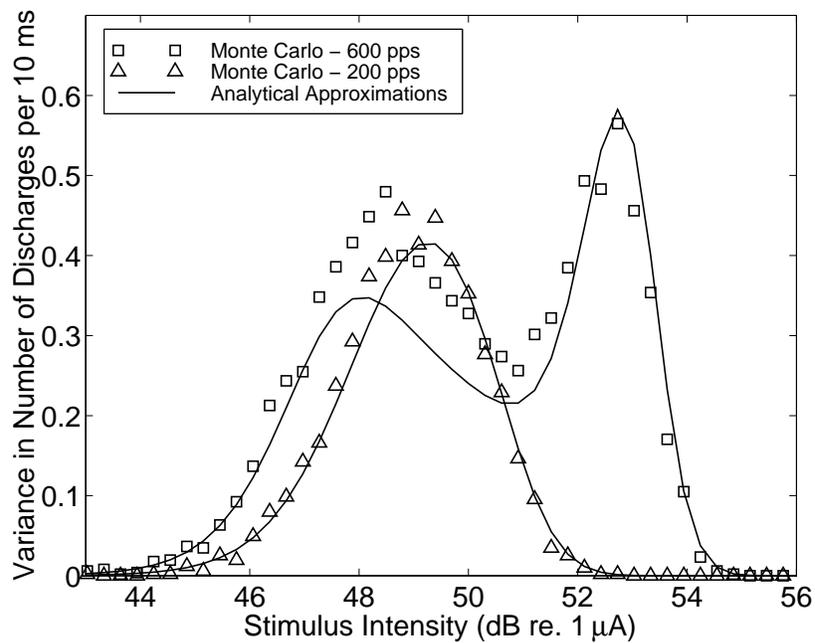
In order to test the accuracy of this analytical approximation, Monte Carlo simulations of the stochastic model were conducted. Responses to 10, 20 and 100 ms pulse trains of rates 200 pps and 600 pps were simulated and compared with the results of the analytical approximation. Estimates of discharge rate mean and variance computed by 500 iterations of the simulations and by the analytical approximation for 10, 20 and 100 ms pulse trains are plotted in Figures 7-3, 7-4 and 7-5 respectively.

For a 10 ms pulse train, the analytical description approximates the response at 600 pps reasonably accurately and the response at 200 pps very well. Extending the pulse train duration to just 20 ms (4 pulses at 200 pps and 12 pulses at 600 pps) improves the accuracy significantly. For a 100 ms pulse train, the analytical expressions provide an almost exact description of the model. This indicates that the accuracy of the analytical descriptions is relatively insensitive to the length of the pulse train, i.e., to the accuracy of Approximation 7.3.2.

Furthermore, the analytical approximation takes just 0.04% of the time required by the Monte Carlo simulation to compute 500 iterations for a 100 ms pulse train. While the mean of the Monte Carlo stimulation output converges rapidly, giving an accurate estimate after only 100 iterations,

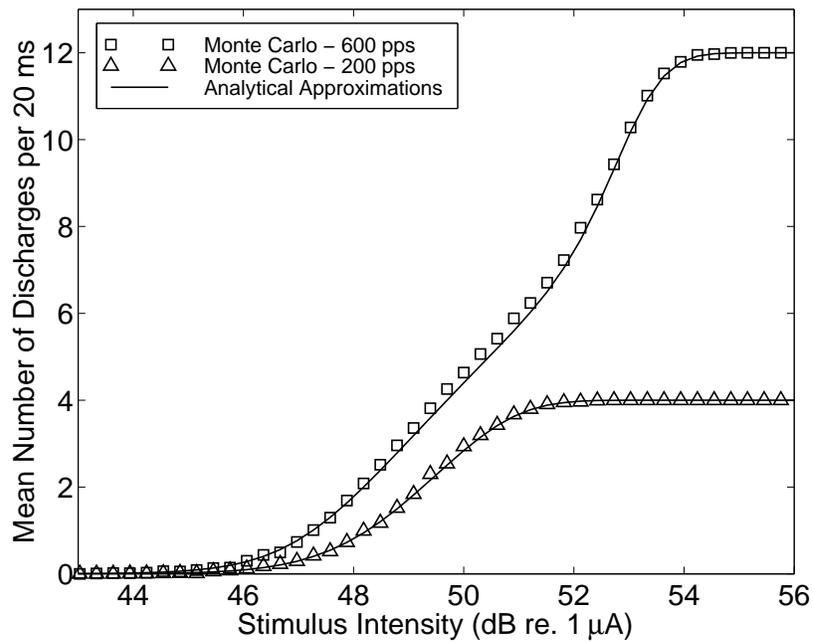


(a) Mean discharge rate versus stimulus intensity.

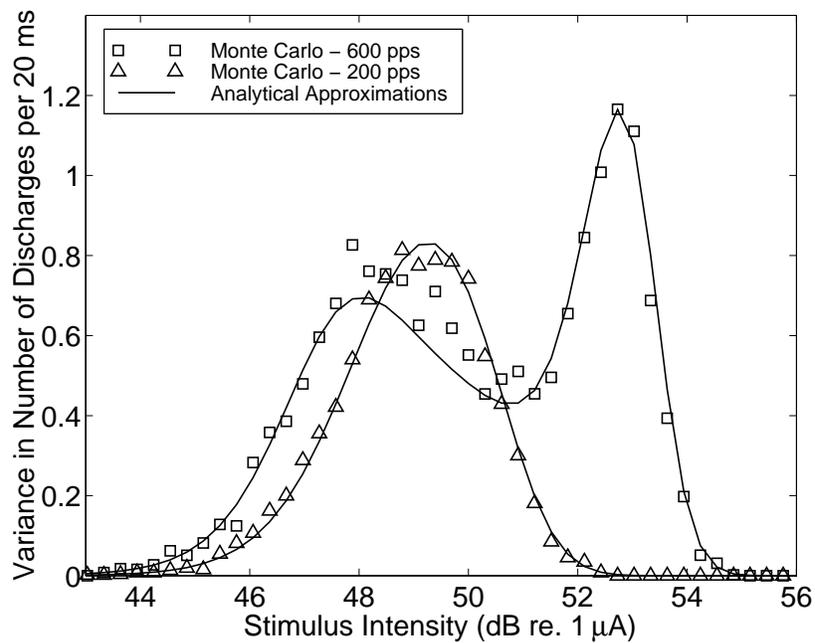


(b) Discharge rate variance versus stimulus intensity.

Figure 7-3: 10 ms pulse train: Comparison of discharge rate mean and variance predicted by 500 iterations of Monte Carlo simulations for pulse rates of 200 pps ( $\triangle$ ) and 600 pps ( $\square$ ) and by the analytical approximation (solid lines).

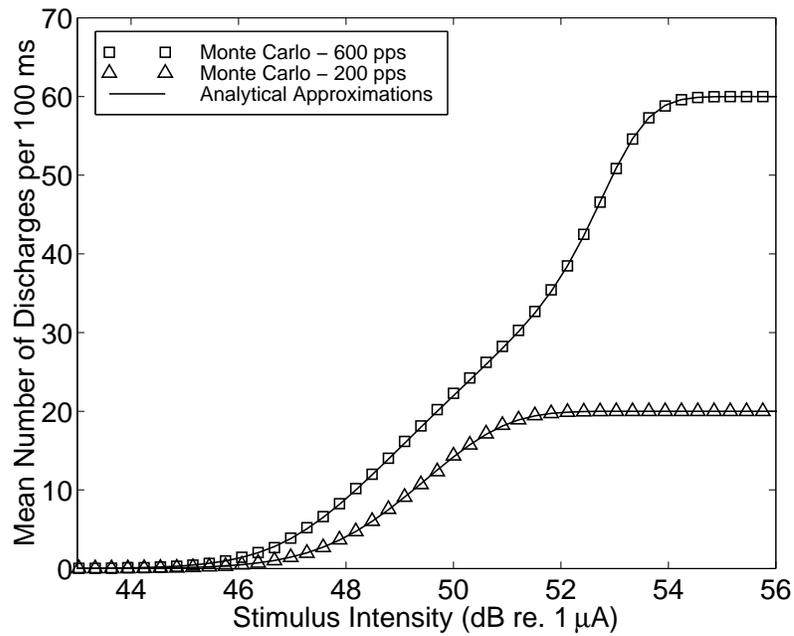


(a) Mean discharge rate versus stimulus intensity.

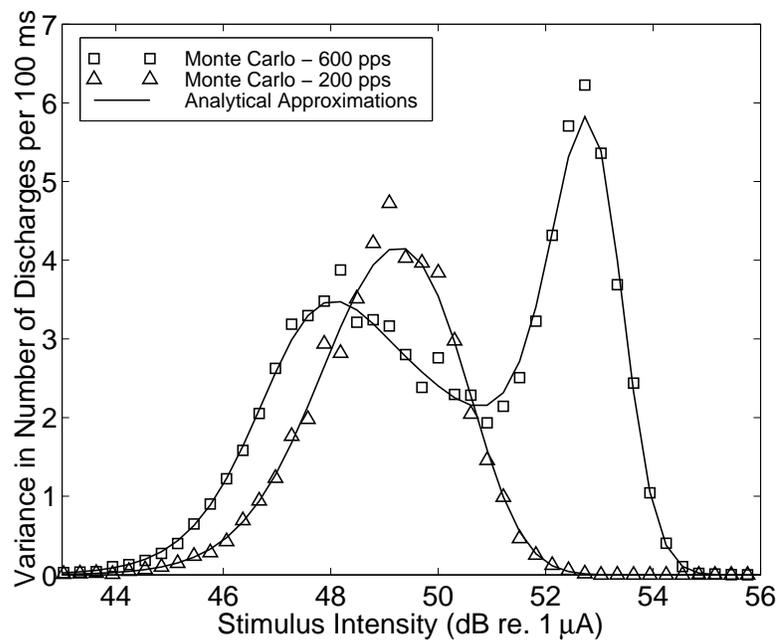


(b) Discharge rate variance versus stimulus intensity.

Figure 7-4: 20 ms pulse train: Comparison of discharge rate mean and variance predicted by 500 iterations of Monte Carlo simulations for pulse rates of 200 pps ( $\triangle$ ) and 600 pps ( $\square$ ) and by the analytical approximation (solid lines).



(a) Mean discharge rate versus stimulus intensity.



(b) Discharge rate variance versus stimulus intensity.

Figure 7-5: 100 ms pulse train: Comparison of discharge rate mean and variance predicted by 500 iterations of Monte Carlo simulations for pulse rates of 200 pps ( $\triangle$ ) and 600 pps ( $\square$ ) and by the analytical approximation (solid lines).

the variance has still not converged completely after 500 iterations. The sample mean from the Monte Carlo simulations will converge at a rate of  $\sqrt{N}$ , where  $N$  is the number of iterations, and the sample variance will converge more slowly, at a rate of approximately  $\sqrt{(N+1)/2}$  (Keeping, 1962, pp. 108). To double the accuracy of the estimate of the variance would take 7 times the number of iterations, such that the analytical approximation, which does not require multiple iterations, would take just 0.006% of the computational time.

Therefore, no further results of the Monte Carlo simulations are presented here. Instead, predictions of the physiological pulse-train data by the analytical approximation are presented in the next section.

**Remark 7.3.3** An expression for the probability distribution of renewal times, or ISI histogram, can also be directly calculated via the equation:

$$f(n) = \sum_j f(n|j)v_j \quad (7.19)$$

where  $f(n)$  is defined as the probability that the next discharge will occur in pulse  $n$ , given that the last discharge occurred  $n$  pulses previously.

**Remark 7.3.4** The modification of the threshold by the refractory function means that contrary to Approximation 5.3.2, the threshold with which  $V_{\text{stim}}$  is compared *will* change during pulse  $n$  and, depending on the time-course of the refractory function, the time-course of  $V_{\text{stim}}$  may become significant. From this, a case could be made for utilising a stimulus potential which rises monotonically over the duration of the pulse to its asymptotic value. As long as the refractory function follows Assumption 7.3.1, i.e., is monotonically decreasing, only one adjustment would need to be made to the preceding method of finding the mean and variance of the discharge rate. Substituting a monotonically increasing stimulus potential  $V_{\text{stim}}(n_i|j)$  for  $V_{\text{stim}}(n|j)$  in calculating  $g(n_i|j)$  in (7.4) and (7.5) will change the value of  $p(n_i|j)$  as calculated via (7.12), and in turn the value of  $B_{ij}$  as calculated via (7.11). This could give a more accurate description of the probability of an action potential occurring in a particular bin of a pulse, and consequently could give a more accurate prediction of the mean and variance of the renewal time.

### 7.3.4 Prediction of pulse-train response

Following the method derived in the previous section, both the mean and variance of the discharge rate can be predicted for the pulse train model when RS is equal to zero (Deterministic Model), as well as for *Case 1* and *Case 2* as defined in Section 7.3.1. Taking the values for threshold and RS

from fits to the single-pulse response done in Chapter 5, it simply remains to choose an appropriate refractory function.

The single-pulse response was calculated by investigating the response of each neuron to the first pulse in a pulse train. It is therefore possible to directly calculate the refractory function which produces the best fit to the discharge rate data for the remainder of the pulse train. However, discharge rate data at a range of stimulation rates were collected for only a few neurons in this data set. Consequently, it is infeasible to estimate the statistics of refractory functions in AN fibres, for use in the total AN model. Therefore, a standard refractory function is used for all neurons which is derived by fitting an exponential function to the cat data from Figure 3-5 of Dynes (1996). These data were collected using monophasic pulses, but are consistent with biphasic pulse data (e.g., see Fig. 20 of Parkins, 1989). The standard refractory function is plotted in Figure 7-6 as the relative increase in threshold, i.e.,  $(V_{\text{thr}} + V_{\text{refr}}) / V_{\text{thr}}$ , along with its analytical description. Note that the absolute refractory period lasts 0.7 ms, after which the relative refractory period begins with a time-constant of 1.32 ms, with the refractory effects lasting for a total period of 20 ms. The effect of refractory function choice is investigated in the next section.

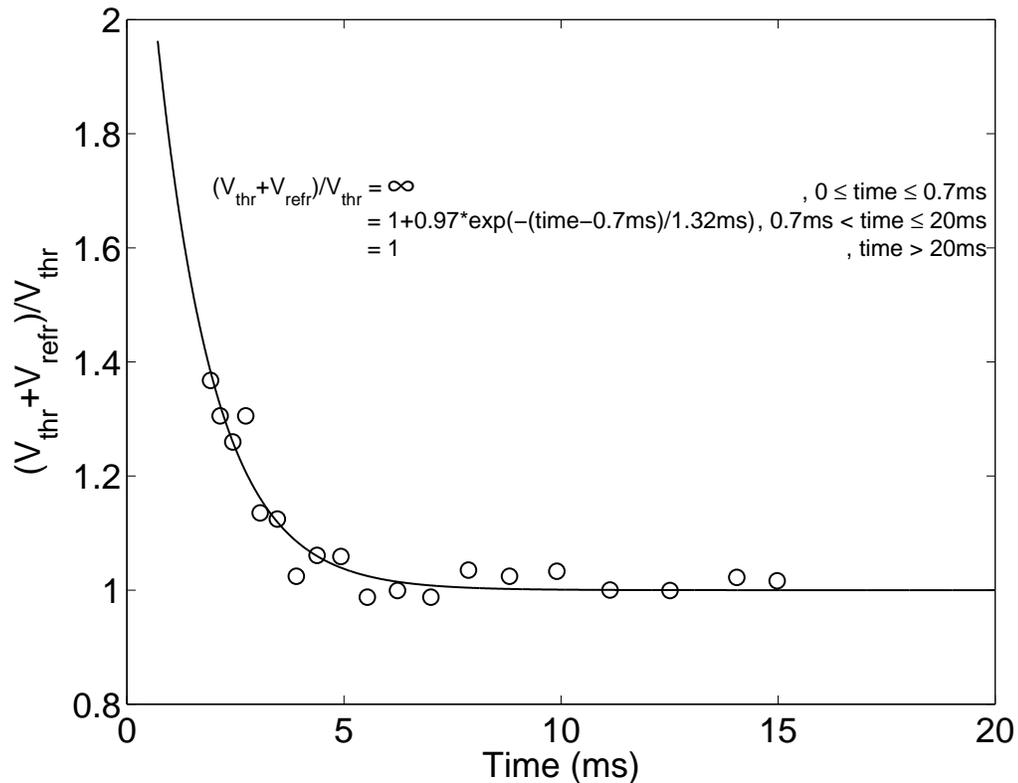


Figure 7-6: Relative increase in threshold due to refractory effects: physiological data (o) and exponential least-squares fit (solid line).

### Mean discharge rate

Plotted in Figure 7-7 are mean discharge rates for Neuron 3-21 of Javel et al. (1987) at pulse rates of 100, 200, 300, 400, 600 and 800 pps with a pulse duration (pulse width) of  $100 \mu\text{s}/\text{phase}$ . There were a good deal of responses at the higher stimulus levels for which no data were collected. As the stimulus amplitude was increased, the neurophysiologists would stop collecting data once the discharge rate equalled the stimulus's pulse rate. This was done because they had previously found that the spike rate would remain nearly constant and equal to the pulse rate at these higher stimulus levels. The resulting extrapolations are plotted as dashed lines.

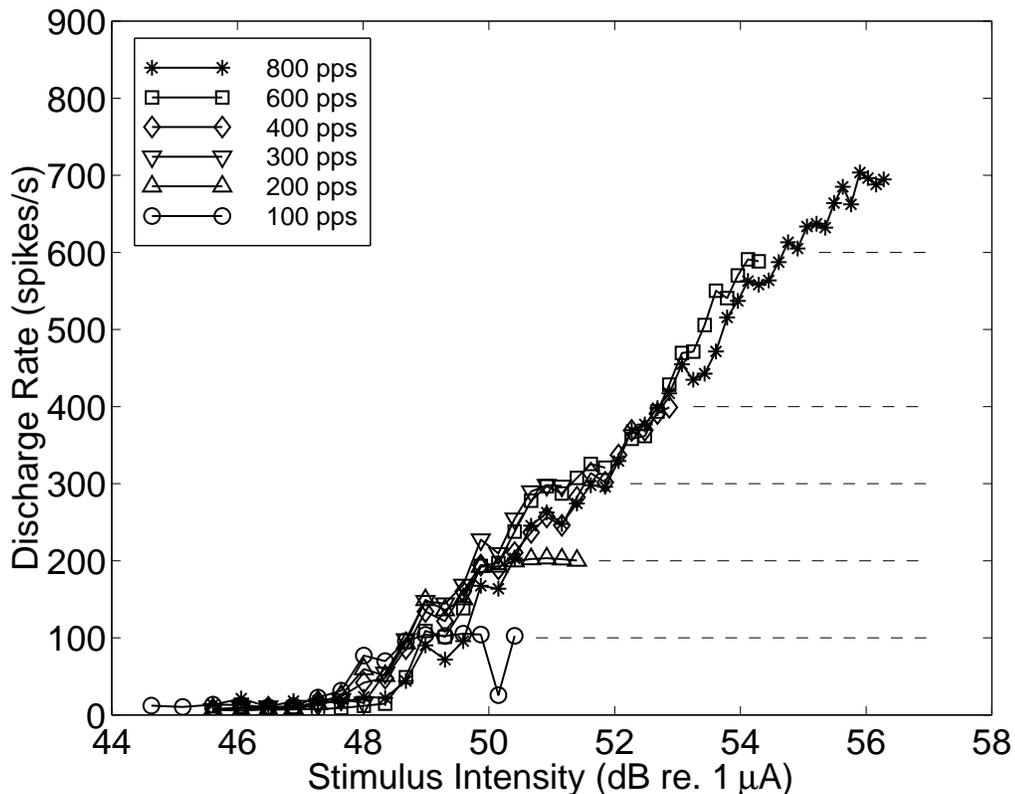


Figure 7-7: Mean discharge rates for Neuron 3-21 of Javel et al. (1987) at pulse rates from 100 to 800 pps with a pulse width of  $100 \mu\text{s}/\text{phase}$ . Dashed lines indicate assumed extrapolations of the data.

The data from Figure 7-7 are replotted in Figure 7-8, along with the first-pulse data of Neuron 3-21 (dotted lines) scaled to indicate what the pulse-train response would be if inter-pulse interactions were not present. This comparison shows that the slopes of the I/O functions in response to pulse trains are flattened due to inter-pulse interactions.

Plotted in Figure 7-9 are the deterministic pulse-train model predictions of these data (solid lines). The dotted lines represent the same predictions without refractory effects, obtained from the deterministic single-pulse model developed in Chapter 5. When compared with the data in

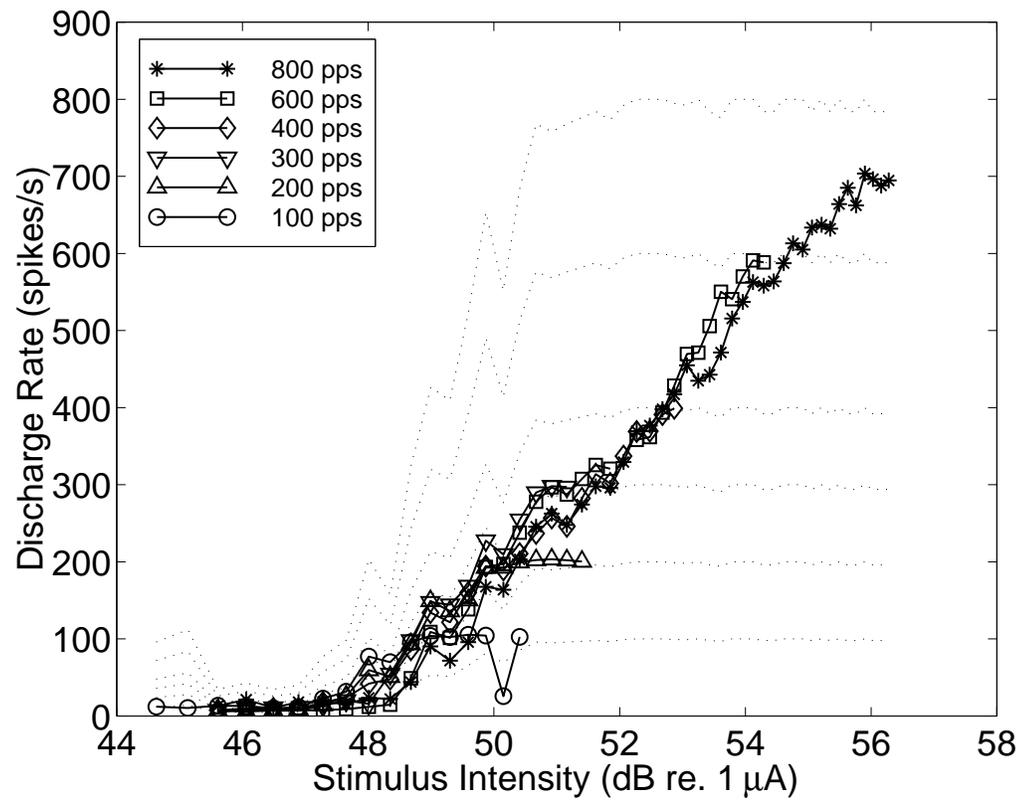


Figure 7-8: Mean discharge rates for Neuron 3-21 of Javel et al. (1987) at pulse rates from 100 to 800 pps with a pulse width of  $100 \mu$ s/phase. Dotted lines are the first-pulse data of Neuron 3-21 scaled to indicate what the pulse-train response would be if inter-pulse interactions were not present.

Figure 7-7, the inclusion of the refractory mechanism is seen to improve the model's prediction of the I/O function slope for each pulse rate, but it still does not reach the slope of the physiological data and the shapes of the curves are inaccurate.

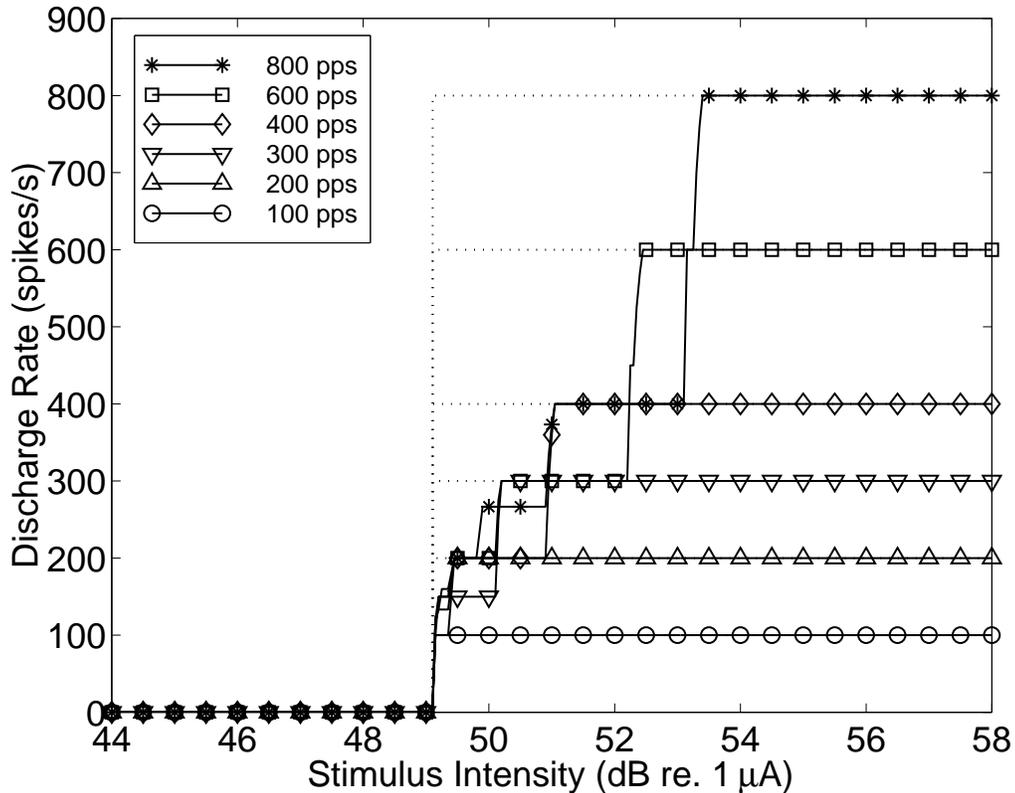


Figure 7-9: Mean discharge rates for the deterministic version of the model at pulse rates of 100, 200, 300, 400, 600 and 800 pps with a pulse width of  $100 \mu\text{s}/\text{phase}$ . Dotted lines indicate mean discharge rates without refractory effects.

Plotted in Figure 7-10 are predictions of the mean discharge rates for the *Case 1* version of the stochastic pulse-train model. The dotted lines represent the same predictions without refractory effects, obtained from the stochastic single-pulse model developed in Chapter 5. Both the shape of the curves and their slopes better predict the physiological data than the deterministic model does.

The effects of the choice of refractory function have also been investigated. Plotted in Figure 7-11 are mean discharge rates where the refractory function has been modified to have an absolute refractory period of 1 ms and a time-constant of 2 ms. For comparison, the discharge rate functions for the standard refractory function from Figure 7-10 (dashed lines) have been overlaid. The modified refractory function produces a slightly better prediction of the overlap of the physiological discharge rate versus intensity curves at discharge rates below approximately 200 spikes/s than the standard refractory function does. However, the shapes of the curves and the slopes appear to

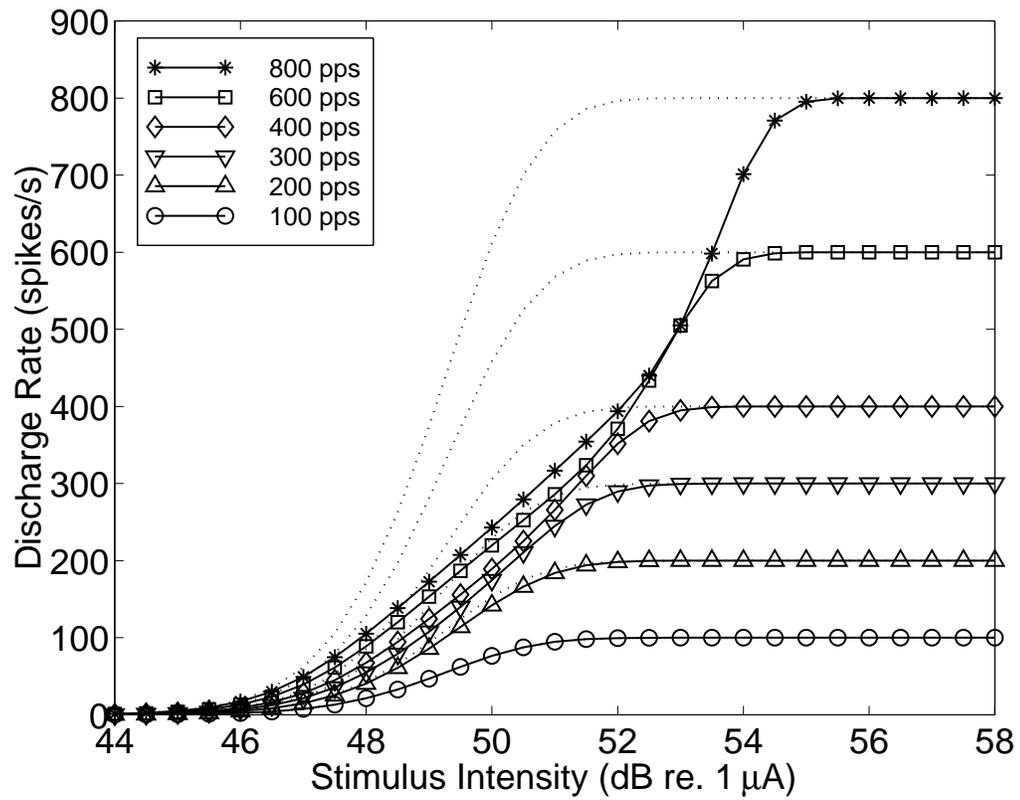


Figure 7-10: Mean discharge rates for the standard version of the stochastic model (i.e., *Case 1*) at pulse rates of 100, 200, 300, 400, 600 and 800 pps with a pulse width of 100  $\mu$ s/phase. Dotted lines indicate mean discharge rates without refractory effects.

be less accurate.

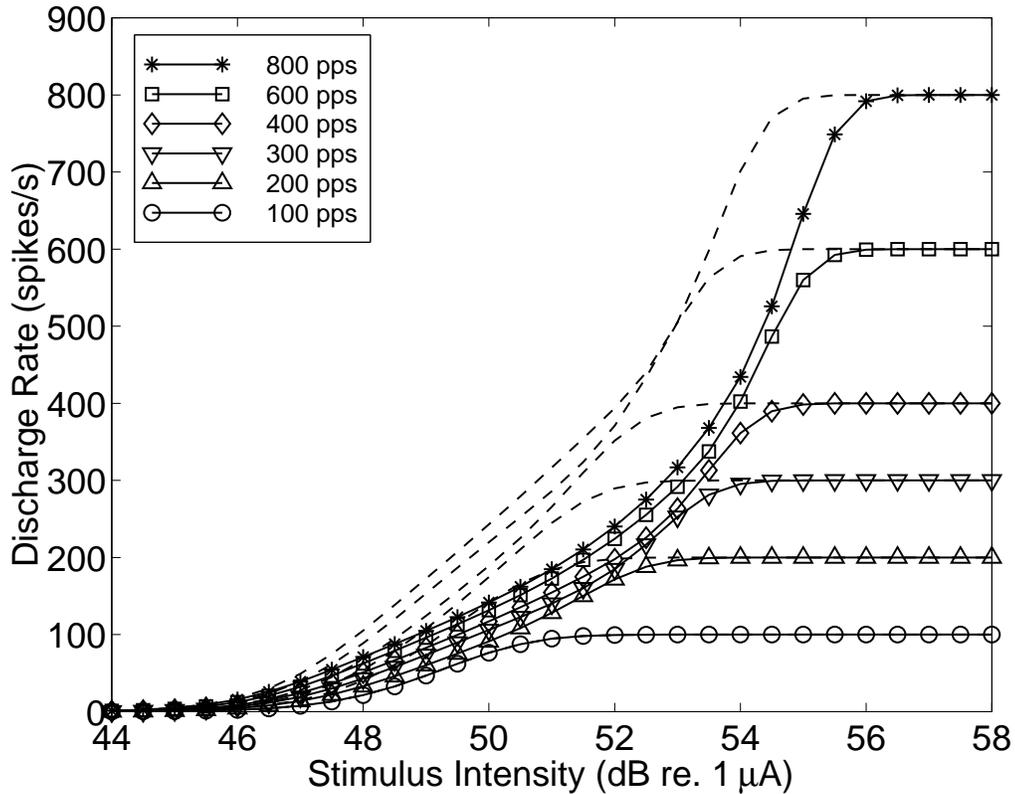


Figure 7-11: Mean discharge rates for the stochastic model with a modified refractory function (absolute refractory period = 1 ms; time-constant = 2 ms) at pulse rates of 100, 200, 300, 400, 600 and 800 pps with a pulse width of 100  $\mu$ s/phase. Dashed lines indicate mean discharge rates for the standard refractory function (from Figure 7-10).

Plotted in Figure 7-12 are mean discharge rates for the *Case 2* version of the stochastic model. Dashed lines indicate mean discharge rates for the *Case 1* version. The *Case 2* version of the model may provide a slightly better prediction of the shapes and slopes of the physiological discharge rate versus intensity curves than the *Case 1* version does.

Mean discharge rate data at stimulation rates of 100, 200, 300 and 400 pps from neurons 3-10 (see Figure 7-13) and 3-29 of Javel et al. (1987) were also investigated, yielding very similar results.

### 7.3.5 Variance of discharge rate

In Figure 7-14 discharge rate variances ( $\sigma$ ) are plotted for Neuron 3-21 of Javel et al. (1987) at pulse rates of 200 and 600 pps. To allow comparison across different pulse rates, both the mean and variance discharge-rate data are normalised (divided by the pulse rate). This normalisation produces the probability of discharge per pulse and the variance per pulse respectively. It can be seen that with increasing pulse rate the variance is compressed around the point of 0.5 probability

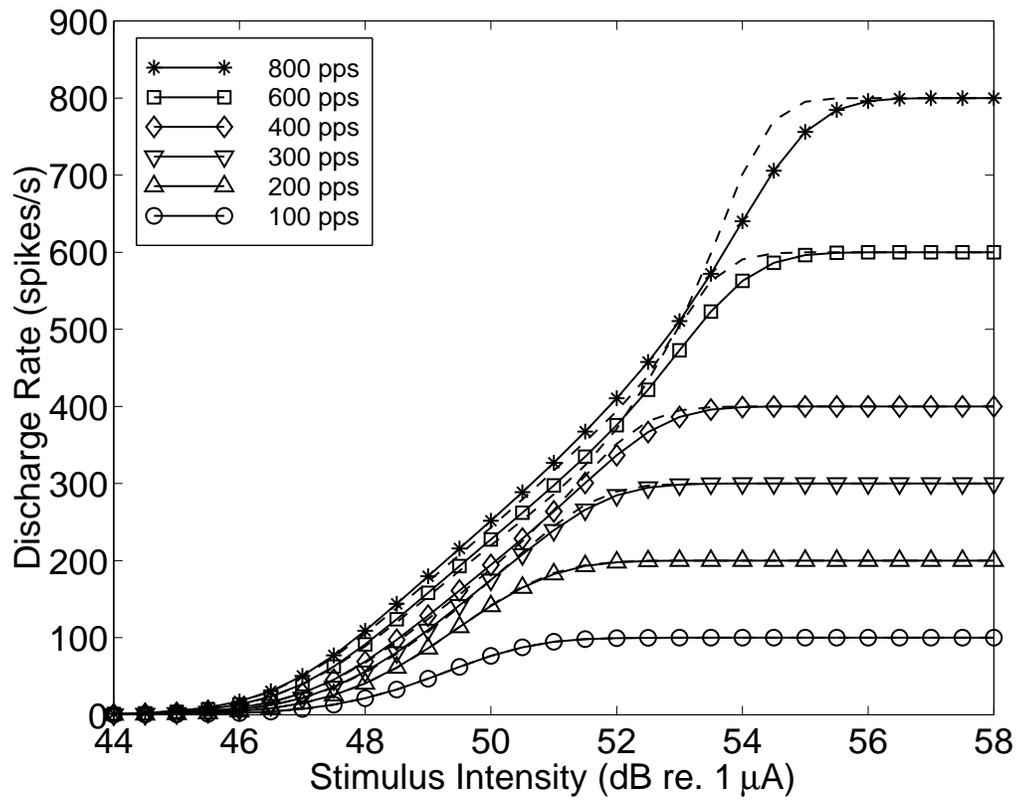
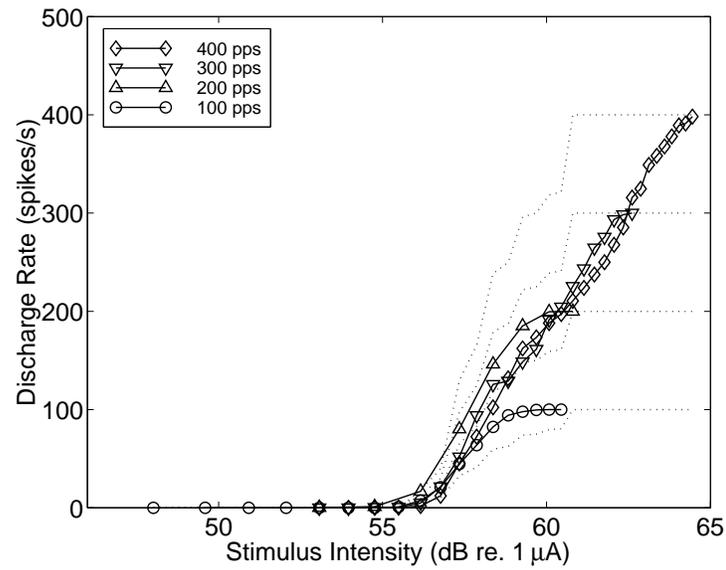
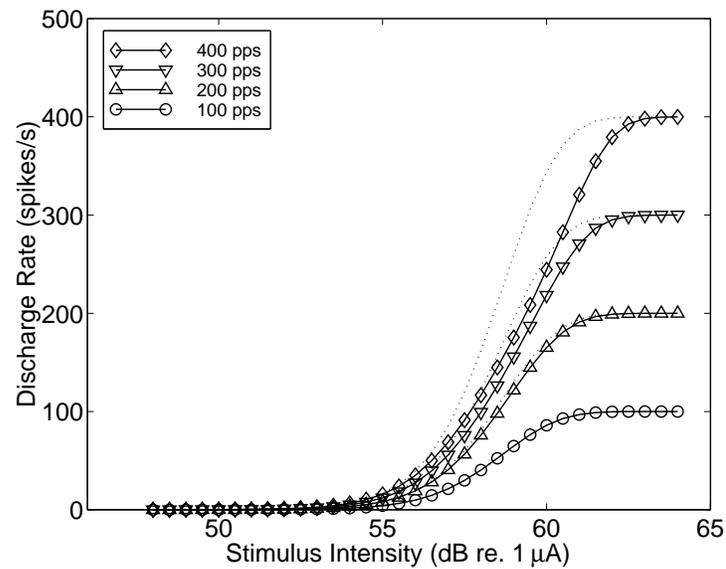


Figure 7-12: Mean discharge rates for the *Case 2* version of the stochastic model at pulse rates of 100, 200, 300, 400, 600 and 800 pps with a pulse width of 100  $\mu\text{s}$ /phase. Dashed lines indicate mean discharge rates for the *Case 1* version (from Figure 7-10).



(a) Mean discharge rate data (solid lines) from Neuron 3-10. Dotted lines are the first-pulse data of Neuron 3-10 scaled to indicate what the pulse-train response would be if inter-pulse interactions were not present.



(b) Stochastic model predictions (solid lines) of mean discharge rate data from Neuron 3-10. Dotted lines indicate mean discharge rates without refractory effects.

Figure 7-13: Mean discharge rate data and model predictions for Neuron 3-10 of Javel et al. (1987) at pulse rates from 100 to 400 pps with a pulse width of 200  $\mu\text{s}$ /phase.

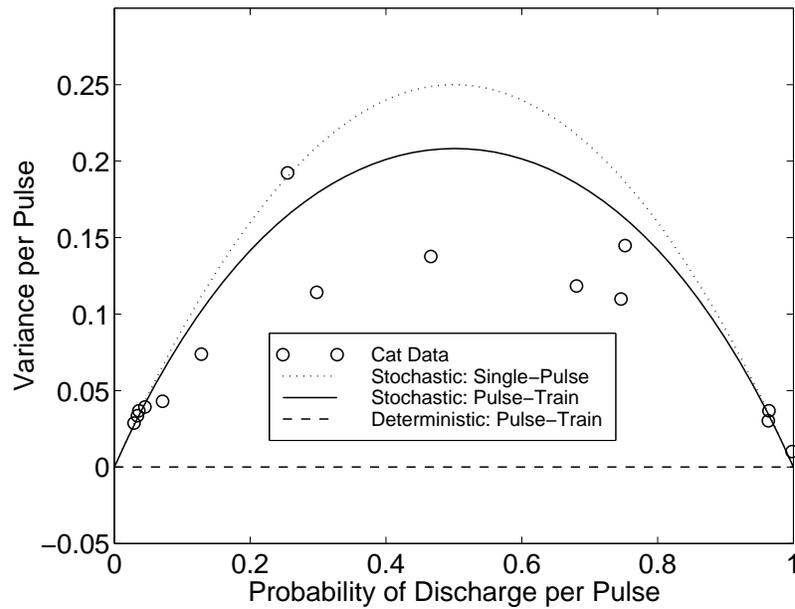
of discharge per pulse. The single-pulse version of the stochastic model (dotted line) overestimates the variance, the error increasing with pulse rate. The deterministic model (dashed line) on the other hand cannot predict the variance at all. The *Case 1* version of the pulse-train stochastic model (solid line) provides the best prediction of the variance data. This was found to be true also for pulse rates of 100, 300, 400 and 800 pps.

Just as the effects of refractory function choice and of *Case 1* versus *Case 2* on the mean discharge rate were investigated, their effect on the prediction of the variance data are also investigated. In Figure 7-15 the discharge rate variance data (o) are plotted again, along with the *Case 1* model predictions with the standard refractory function (dotted lines). Also plotted are the predictions of the stochastic model (solid lines) with a modified refractory function (absolute refractory period = 1 ms; time-constant = 2 ms) and the predictions of the *Case 2* version of the stochastic model (dashed lines). It can be seen that the change in refractory function time-constants improves the model predictions of the variance data more than the change from *Case 1* to *Case 2* does. These improvements are seen to occur specifically in the range from 0.1 to 0.6 probability of discharge per pulse.

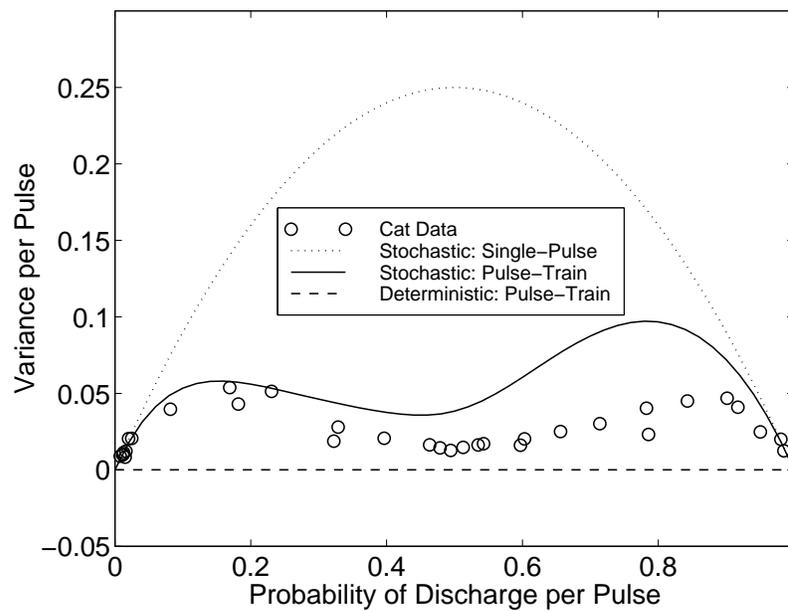
Variance of discharge rate data at stimulation rates of 200 and 400 pps from neurons 3-10 (see Figure 7-16) and 3-29 of Javel et al. (1987) were also investigated, yielding very similar results.

## 7.4 Total auditory nerve response

The output of the model of total auditory nerve response developed in Chapter 5 is the mean and variance in discharge rate for each neuron. Although for responses to pulse trains the individual neurons are described by renewal processes rather than Bernoulli processes, as was the case for the single-pulse model, application of the Central Limit Theorem (Devore, 1987) shows that the total response will still be well approximated by a Gaussian distribution if the mean is large enough ( $\geq 15$ ). If the mean is less than 15, the Poisson distribution approximation of the total response will worsen if one or more fibres diverge from behaving like a Bernoulli process. As is evident from Figures 7-14 and 7-15, the renewal process model of a fibre's behaviour (stochastic pulse-train model) differs from a Bernoulli process (stochastic single-pulse model) only when its discharge probability per pulse becomes larger than about 0.05. For total cochlear spike counts of less than 15, fibre discharge probabilities above 0.05 could only occur: (i) if the density of surviving AN fibres is very low, or (ii) if the stimulating electrode is very selective and therefore effectively stimulates only a small number of fibres.

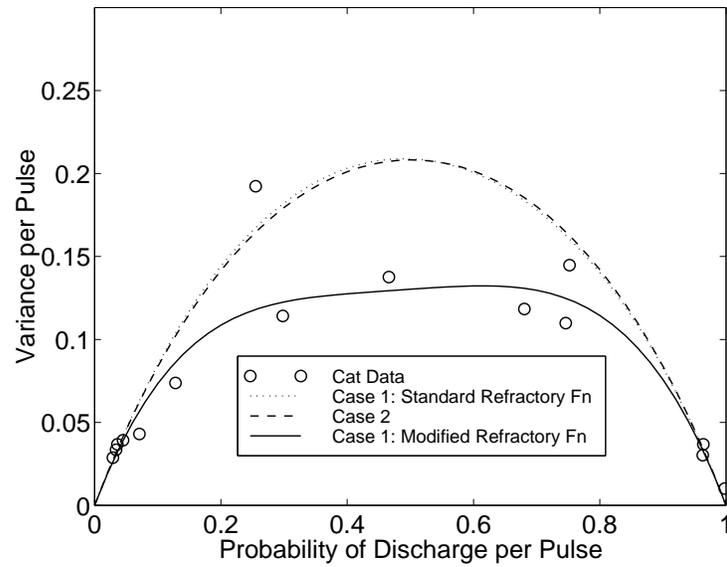


(a) Variance per pulse at 200 pps

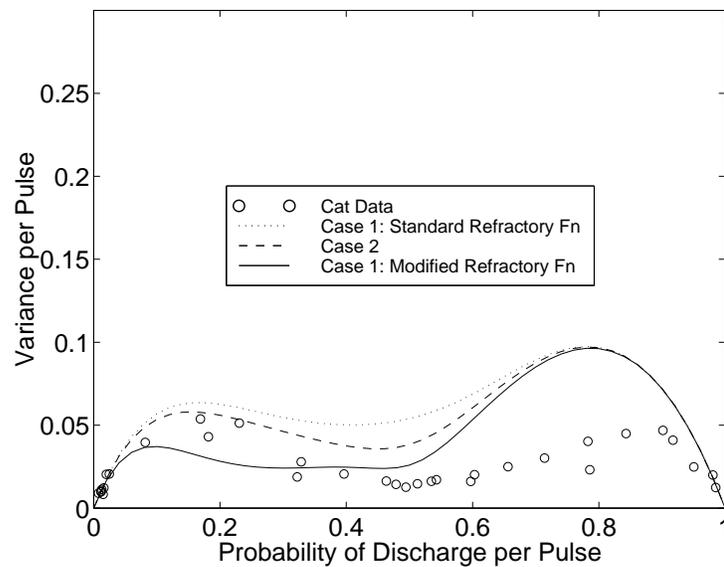


(b) Variance per pulse at 600 pps

Figure 7-14: Variance per pulse versus probability of discharge per pulse (o) for Neuron 3-21 of Javel et al. (1987) with a pulse width of  $100 \mu\text{s}/\text{phase}$ . Also plotted are the single-pulse stochastic model predictions (dotted lines) of these data, the stochastic model (*Case 1* version) predictions (solid lines) and the deterministic model predictions (dashed lines).

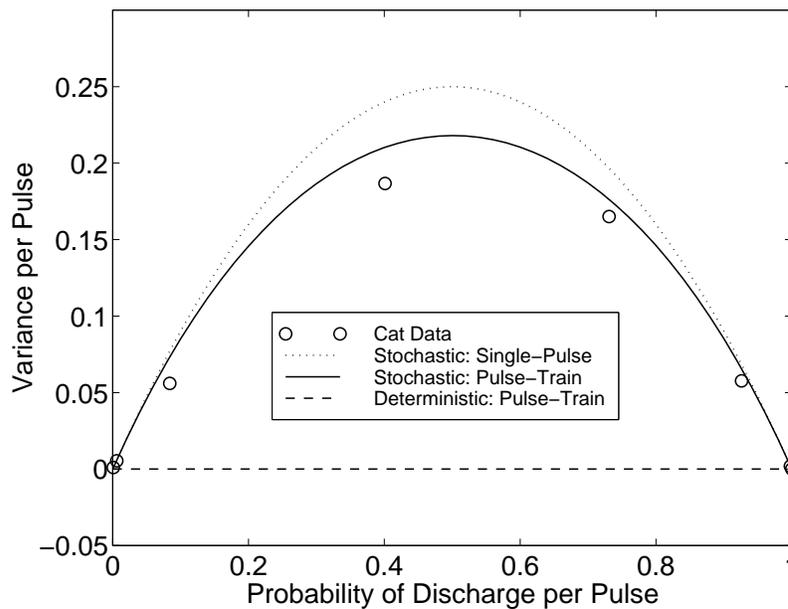


(a) Variance per pulse at 200 pps

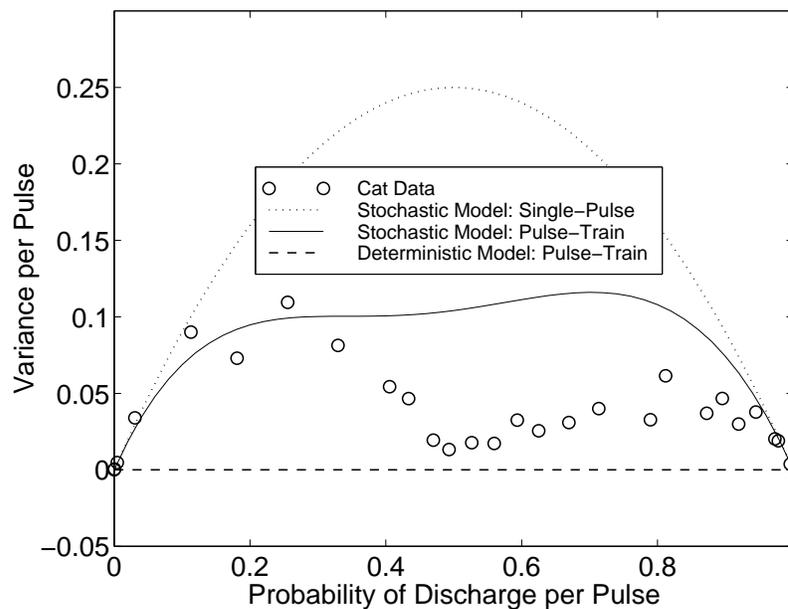


(b) Variance per pulse at 600 pps

Figure 7-15: Variance data (o) from Figure 7-14 are replotted, as are the stochastic model predictions of these data with the standard refractory function (dotted lines). Also plotted are the *Case 2* model predictions (dashed lines) and the stochastic model predictions with a modified refractory function (solid lines).



(a) Variance per pulse at 200 pps



(b) Variance per pulse at 400 pps

Figure 7-16: Variance per pulse versus probability of discharge per pulse (o) for Neuron 3-10 of Javel et al. (1987) with a pulse width of  $100 \mu\text{s}/\text{phase}$ . Also plotted are the single-pulse stochastic model predictions (dotted lines) of these data, the stochastic model (*Case 1* version) predictions (solid lines) and the deterministic model predictions (dashed lines).

## 7.5 Discussion

The results of this chapter show that a deterministic model is a poor description of a neuron's response, in light of the improvements in the prediction of physiological data when a stochastic component of response is added to the neural model. Specifically, the results show that a simple stochastic model can accurately predict the mean and variance of the discharge rate in response to pulse trains. The results suggest that in the case where a uniform refractory function is assumed across all neurons, just two parameters (threshold and RS) can well describe the response of an auditory nerve fibre to single biphasic pulses and pulse trains. The results show that an even more precise description may be achieved by estimating a specific refractory function for each neuron. However, the improved accuracy gained by estimating refractory functions for each neuron may not justify the additional parameters required to describe each neuron in the total auditory nerve model.

The analytical descriptions of single fibre response to pulse trains developed in this chapter provide a simple and computationally efficient method of modelling the response of a large-scale population of AN fibres to trains of evenly-timed, uniform pulses. Furthermore, the output of the model, being a mean and a variance for each neuron with a known distribution, allows the direct use of signal detection theory to determine the resolution with which the AN can convey information about an electrical stimulus.

The analytical approximation was developed to describe responses to trains of evenly-timed, uniform pulses, as are commonly used in psychophysical and physiological investigations. However, current speech processing strategies used in cochlear implants may produce trains of pulses which vary in amplitude over the duration of the train. Consequently, rather than the stationary process produced by a uniform pulse-train, neural response to non-uniform pulse trains would need to be modelled by a *nonstationary* process. Neither the analytical descriptions, nor the methods used in deriving them, can be directly applied to nonstationary processes. Analytical expressions for nonstationary processes may also be tractable, but if not, Monte Carlo simulations of the model illustrated in Figure 7-1 may be used to estimate the neural response.

Further physiological and modelling studies would be helpful in refining and extending a number of aspects of the model.

In this chapter randomness in the discharge rate in response to electrical stimuli has been investigated. In developing an analytical description of discharge rate, the model's description of the timing of the discharges has been degraded. As was the case in Chapter 5, a separate stochastic description of action potential latency would be needed to investigate the effects of jitter in theories of coding which are dependent on the exact timing of discharges.

Discharge rate data at a range of stimulation rates were only available for three neurons in the cat data analysed in this chapter. It would be desirable to compare the model with more

physiological data, for further verification of the model. Additional data could also help in further refining the model. In particular, an estimate of the distribution of refractory functions in the cat AN could be obtained if more data were available.

The stimulation rates investigated in this chapter were all 800 pps or lower. Development of speech processing strategies utilising significantly higher stimulation rates is currently being investigated by a number of cochlear implant research groups. It would therefore be useful to extend the model to describe physiological data collected at these higher stimulation rates when such data become available.

Furthermore, much of the neurophysiology and, as a consequence, much of the modelling in this chapter emphasises moderate to high discharge probabilities per pulse. Implant users may be operating at considerably lower discharge probabilities, particularly when receiving continuous stimulation from a multi-channel speech processor (see Chapter 6). It may therefore be useful to collect further data at stimulus intensities which produce low discharge probabilities per pulse.

Further investigation into the behaviour of membrane noise following both subthreshold and suprathreshold stimulation is required. Single-fibre data from the cat AN and one stochastic model (see Figure 5-7 of Dynes, 1996) indicate that the standard version of the stochastic model (*Case 1*) is the more likely description of membrane noise following a suprathreshold pulse, i.e., the membrane noise does not change significantly during the refractory period. However, membrane noise was seen to increase following a subthreshold pulse, i.e., in response to a depolarisation which does not generate an action potential. This may explain why, for mid to high pulse rates, the variance per pulse in Javel et al.'s (1987) cat data is slightly greater for probabilities of discharge per pulse below 0.5 than for discharge probabilities above 0.5 (e.g., see Figure 7.16(b)): for probabilities of discharge per pulse below 0.5, the majority of the pulses in a pulse train are effectively subthreshold, producing greater membrane noise and consequently greater variance in the response. This phenomenon could be very significant if cochlear implants are operating at intensities which produce low discharge probabilities, such that many pulses in a pulse train will effectively be subthreshold.

A different stochastic model (Rubinstein, Matsuoka, Abbas and Miller, 1997) showed that in some cases membrane noise may increase significantly during a portion of the refractory period. Single-fibre data from the cat AN (Miller et al., 1997) also revealed that some neurons exhibit double-spiking and that the stochastic component of the second spike is significantly increased. This may be a result of the membrane noise increasing by such an amount during the refractory period following the first spike that a second spike is spontaneously generated (Miller et al., 1997). These phenomena could also be significant for cochlear implant speech processing strategies and therefore further investigation of these phenomena is desirable.

As well as choice of refractory function and the effects of subthreshold and suprathreshold

stimulation on noise levels, the other inter-pulse interactions discussed in Section 7.3.1, but not included in this model, may account for some of the differences seen between the variance data at mid to high discharge rates and the model's predictions of these data (e.g., see Figures 7-14 and 7-16). However, the efficiency of characterising nonlinearities using uniform, evenly-timed pulse trains may be limited—inter-pulse interactions may be better investigated using non-uniform pulse trains as used in engineering methods of nonlinear system identification (e.g., Shi and Hecox, 1991).

As was the case for the single-pulse model, the physiological data presented in this chapter and the models derived from this data were from cats. Care should be taken in extrapolating the model to other species, including human cochlear implant users. Furthermore, the physiological data are from normal hearing cats. Other than assuming zero hair cell survival and allowing an arbitrary number of surviving AN fibres, this model does not take into account the effects which prolonged deafness has on the response of AN fibres to electrical stimulation (Shepherd and Javel, 1997). An extension to this study could be to model the effects of various aetiologies on single fibre I/O and refractory functions.

In Chapter 6 the stochastic single-pulse model of Chapter 5 was used to predict a range of psychophysical measures. The results of that study suggest that cochlear implants are indeed operating in the region where stochastic activity is significant. In all the cases investigated, the stochastic model predicted psychophysical performance significantly better than the deterministic model did. This suggests that the stochastic single-pulse model is not just more accurate in its prediction of physiological response, but also in its prediction of the resulting behavioural performance. With the pulse-train model developed in this chapter, an appropriate tool is now available for extending this investigation to the prediction of psychophysical data for higher stimulation rates. The pulse-train model could also be used in psychophysical models of temporal frequency coding, such as those investigated in Chapters 2 to 4, to investigate how the pitch percept related to the stimulation rate may arise.

# Chapter 8

## Conclusions

### 8.1 Summary

It is known that in a single fibre the predominant interval between action potentials is the same as the period of a pure-tone acoustical stimulus for frequencies up to around 600 Hz, and this is a possible mechanism by which the brain codes frequency. For higher frequencies, refractory effects prohibit a single fibre from firing every cycle, such that the predominant peak in the interspike interval histogram becomes a multiple of the period of the stimulus. However, the difference between the interspike intervals in single fibres represents the period of the stimulus for frequencies up to 1.5–3.0 kHz. This suggests that interspike intervals in groups of fibres could be used for coding frequency in this higher frequency range. In Chapter 2 interspike intervals across a pair of auditory nerve fibres were investigated to determine how they may code the stimulus frequency. Computer models of the cochlea and the hair-cell/auditory-nerve transduction were used to investigate the cross-fibre interspike interval probability distribution in different areas in the cochlea. The response to pure tones was simulated by the composite model which provides a spatiotemporal excitation pattern representing the fluctuating firing probability of the auditory nerve fibres. The cross-fibre interspike interval probability distribution for fibres innervating different areas in the cochlea was then calculated from their temporal firing probabilities using an integral expression. The results of this study showed that temporal frequency coding can take place across fibres, but it can be affected by factors such as the frequency of the stimulus and the phase delay due to the travelling wave. The synchronisation between fibres showed that temporal frequency coding may occur across fibres only for frequencies below 1–3 kHz, if the fibres originate from sites nearby enough in the cochlea for their phase difference to be smaller than 6–8 radians.

In Chapter 3 this investigation was extended to spatiotemporal combination of activity in *more* than two fibres, specifically spatiotemporal summation of auditory nerve firings. Anatomical and physiological studies have found a number of different cell types in the cochlear nucleus which

appear to perform such summations. In Chapter 3 it was determined theoretically how various degrees of auditory nerve fibre convergence onto summing-neurons in the cochlear nucleus would affect estimation and discrimination of stimulus features from the summing-neurons' outputs. To achieve this, an analytical model, based on a point process representation of auditory nerve activity, was derived. This model included parameterised measures of stimulus properties. The summing-neuron model allows for an arbitrary degree of auditory nerve fibre convergence, thus facilitating an investigation of the feasibility of theories of neural coding of sound for different summing-neuron types. Results obtained from the model indicated that for coding of the stimulus frequency in the periodicity of firings could occur for frequencies below 1–3 kHz. To reach such limits, the optimal spread of input fibres is very narrow and temporal integration is very short. This suggests that auditory nerve fibres which converge onto cochlear nucleus cells exhibiting *enhanced* synchronisation originate from sites on the basilar membrane very close together. In contrast, cochlear nucleus cell types which exhibit *reduced* synchronisation could be useful in performing estimation and discrimination tasks based on the average discharge rate. The model indicated that this could be achieved by a very wide spread of inputs and/or long-term temporal integration. It was postulated that a wide spread of inputs would be preferential to long-term temporal integration, because this would enable the neuron's output to follow rapid changes in the average discharge rate of a population of fibres. Consequently, transients in the stimulus intensity which produce rapid changes in the average discharge rate of a population of fibres could be detected by a rapid change in the output of such a summing neuron.

In Chapters 2 and 3 models of auditory nerve response to acoustical stimulation were utilised which assumed identical auditory nerve fibres, and only responses to pure-tone stimuli were investigated. However, auditory nerve fibres do differ in their spontaneous (zero input) response, and also in their onset and saturation thresholds. Experiments have shown that fibres with low spontaneous rates show enhanced responses to the envelopes of complex sounds, while fibres with higher spontaneous rates respond to the temporal fine structure. In Chapter 4, an expression was determined for the Cramer–Rao bound for frequency estimation of the envelope and fine structure of complex sounds by groups of fibres with parameterised response properties. The estimation variances were calculated for some typical estimation tasks, and it was demonstrated how, in the examples studied, a combination of low and high threshold fibres may improve the estimation performance of a fictitious 'efficient' observer. Also, combining neurons with different thresholds may improve the estimation performance of neural systems which are based on the detection of dominant interspike times.

In the case of electrical stimulation, spatiotemporal combination of auditory nerve activity has also been postulated to explain why single-fibre physiological data have been unable to account for some important psychophysical phenomena observed in cochlear implant users. This hypothesised

effect of spatiotemporal combination of auditory nerve activity may require that stochastic components of the single-fibre response to electrical stimulation be considered. However, most models of neural response to electrical stimulation, such as the Hodgkin–Huxley equations, are deterministic and do not naturally describe stochastic activity, for which there is significant physiological evidence. For instance, the range of discharge probabilities measured in response to single electrical pulses cannot be explained at all by deterministic models. In Chapter 5 a stochastic model of single-fibre response to single electrical pulses was developed. Both single-fibre and population responses can be calculated for a range of stimuli. Direct comparisons were made between this model and the same model without the stochastic component. The stochastic model accurately predicted probabilities of discharge measured in response to single biphasic pulses. The results showed that this aspect of stochastic activity in auditory nerve response to electrical stimulation can be described by a simple, computationally efficient model.

As was investigated in Chapter 5, most models of auditory nerve response to electrical stimulation are deterministic, despite significant evidence of stochastic activity in the neurophysiological data. Further to this, psychophysical models and analyses of physiological data using entirely deterministic descriptions of the input and the neural response have been unable to accurately predict many important psychophysical phenomena. In order to determine whether or not inclusion of stochastic activity in neural models improves these predictions, in Chapter 6 a model of loudness—including spatiotemporal summation of auditory nerve firings as investigated in Chapter 3—was derived, which can incorporate either the deterministic or the stochastic description of auditory nerve response from Chapter 5. The application of signal detection theory permits direct prediction of how behavioural threshold, dynamic range and intensity difference limens change with stimulus parameters. In order to avoid the complication of inter-pulse interactions and to enable the use of the model developed in Chapter 5, the investigation in this chapter was restricted to low-rate pulse trains. For all parameters investigated, the inclusion of stochastic activity in the model was found to produce more accurate predictions of behavioural performance than when it was not included. These results indicated that stochastic activity in auditory nerve response is significant for loudness perception of low-rate electrical stimulation.

Stochastic activity is not only present in auditory nerve responses to single pulses, but has also been recorded in responses to trains of electrical pulses. For instance, the variance in discharge rate measured in response to regular pulse trains cannot be explained at all by deterministic models. In Chapter 7 the stochastic model of auditory nerve response to electrical stimulation was extended to describe responses to regular pulse-train stimuli. The extended model accurately predicted means and variances of discharge rates measured in response to pulse trains. These results showed that many aspects of stochastic activity in auditory nerve response to electrical stimulation can be described by a simple, computationally efficient model. The pulse-train stochastic model provides

a means for extending the investigation carried out in Chapter 6 to higher pulse rates, and provides a more precise prediction of the spatiotemporal patterns of response produced by cochlear implant speech processing strategies.

## 8.2 Implications for sound coding by cochlear implants

### 8.2.1 Spatiotemporal combination of auditory nerve activity

In Chapter 2 it was shown how cross-fibre interspike intervals may code stimulus frequencies below 1–3 kHz. In Chapter 3 it was shown how spatiotemporal summation of auditory nerve firings could also be used to code stimulus frequencies over a similar range. These results suggest that using the timing of pulses to code frequency, in addition to the place of stimulation, may be an improvement on *fixed-rate*, spectral-analysis speech processing strategies (Clark, 1997b). One pilot study has investigated the effects of using the timing of pulses to code the envelope of each frequency band in a spectral-maxima speech processor, but no improvement in speech perception was found in the three subjects tested (Vandali et al., 1995). However, further investigation of such schemes is certainly warranted, particularly the introduction of coding of speech features other than the envelope.

The existence of cochlear nucleus neurons receiving inputs for auditory nerve fibres originating from different sites in the cochlea, and therefore having the possibility of performing spatiotemporal summation or cross-fibre interspike interval analysis, may have significant consequences for electrical stimulation via a cochlear implant. It was shown in Chapters 2 and 3 that spatiotemporal combination of auditory nerve activity is highly dependent on the phase relationship of firings in populations of fibres caused by the travelling wave. In the case of electrical stimulation, the phase relationship arising from the basilar membrane vibration is not present, with the consequence that responses to individual pulses will all be very well phase-locked, producing relatively coincident firing in a population of fibres.

Recordings from the anteroventral cochlear nucleus show that neurons which have similar responses to acoustical stimulation can have quite different responses to electrical stimulation (Hocking et al., 1997). Hocking *et al.* found a number of neurons which had primary-like (sustained) responses to acoustical stimulation over a wide range of intensities and frequencies, but only some of these neurons exhibited sustained firing at all stimulus intensities for electrical stimulation rates up to 1 000 pulses per second. Such behaviour could occur if an electrically stimulated neuron receives more input from auditory nerve fibres than it normally would with acoustical stimulation, forcing it into an absolute refractory state which is sustained by continuous input. Cochlear nucleus summing-neurons which receive input from a narrow spread of auditory nerve fibres, as described in Chapter 3, would normally receive very coincident inputs and should not necessarily

receive more input from electrical stimulation than for acoustical stimulation. This may account for neurons which can respond in a sustained manner at all stimulus intensities and rates for both acoustical and electrical stimulation. In contrast, cochlear nucleus summing-neurons which receive input from a wide spread of auditory nerve fibres would receive a staggered input of auditory nerve firings in response to acoustical stimulation because of the cochlear travelling wave. With electrical stimulation this staggering of inputs is not present, so that the amount of coincidence may be increased significantly in response to individual pulses, particularly at high intensities where more fibres respond to each pulse (see Chapter 5). Such behaviour could account for those neurons which cannot respond in a sustained manner at all stimulus intensities and rates. However, from Chapter 6 it appears that cochlear implants may operate at relatively low stimulus intensities under most conditions, such that the effect of ‘increased coincidence’ may be minimal.

Some staggering of pulses on different electrodes *is* used in most speech processing strategies to avoid simultaneous stimulation (Clark, 1987; White et al., 1984; Eddington et al., 1978). One pilot study has shown that speech perception can be improved slightly when using a spectral maxima processing scheme, if the ordering of stimulation on different electrodes follows the direction in which the travelling wave would propagate in a healthy cochlea (Vandali et al., 1995), rather than following the order (ascending or descending) of the maxima’s amplitudes. Implementation of a cochlear model to mimic more precisely the phase relationships produced by the travelling wave (Clark, 1997b; Clark, 1996; Clark, Carter, Maffi and Shepherd, 1995) may improve speech perception further.

In Chapter 6 it was shown that spatial summation of auditory nerve responses to *electrical stimulation* could help in the prediction of loudness perception in cochlear implant users, if the stochastic component of auditory nerve response is considered. Recordings from inferior colliculus neurons in guinea pigs (Montney et al., 1998) show strength-duration curves which match psychophysical data and the stochastic model predictions, unlike the deterministic model predictions. This suggests that the spatial summation described by the psychophysical model is occurring at or before the inferior colliculus, which is consistent with the results of Chapter 3.

### 8.2.2 Effect of neural thresholds on frequency estimation

In Chapter 4 it was shown that auditory nerve fibres with different thresholds are theoretically useful in conveying different components of a complex acoustical stimulus in the periodicity of their firings. In Chapter 5 it was shown that the auditory nerve also exhibits a range of single-fibre thresholds (and input/output function slopes) in response to electrical stimulation. In application to cochlear implants, these results suggest the possibility of encoding different components of complex acoustical stimuli in the firing periodicities of different fibres, by amplitude modulation of the stimulating electrical pulse-train. If the thresholds for electrical stimulation have a strong positive

correlation with those for acoustical stimulation, then amplitude modulation of the pulse train which mimics the amplitude modulation of the basilar membrane vibrations, scaled to compensate for the different range of thresholds and input/output function slopes for electrical and acoustical stimulation, should produce the correct encoding of different components of an acoustical signal in the same auditory nerve fibres as for acoustical stimulation.

However, the factors contributing to single-fibre thresholds in acoustical and electrical stimulation are quite different, and it can be expected therefore that such a strong positive correlation does *not* exist. This means that although it may be possible to encode different components of a complex signal in different auditory nerve fibres, they may not be encoded in the same fibres as for acoustical stimulation. If the mechanisms in the brain used to decode this information are ‘hard-wired’, then this may limit the applicability of amplitude modulation of electrical pulse-trains to convey complex sounds. However, if these mechanisms have some degree of plasticity, they may be able to adapt themselves to the change in which auditory nerve fibres are conveying different components of an amplitude modulated electrical pulse-train.

From Chapter 6 it appears that, because of the broad spread of electrical current generated by a cochlear implant, the saturation level of the input/output function is reached in very few auditory nerve fibres, even at relatively high listening levels. Consequently, not every pulse in a train of evenly-timed, uniform pulses elicits a response from a single fibre. However, amplitude modulation of a pulse-train could be used to lower or raise the discharge probability for particular pulses, without necessarily increasing the average rate of discharge. Under such a scheme the modulation frequency would be well coded by the periodicity of the firings. Furthermore, the present generation of electrode arrays can only be inserted into the *basal* turns of the cochlea, the region of the cochlea which responds to high frequency acoustical stimuli. Fibres innervating this region of the cochlea are unable to maintain synchronisation to their characteristic frequency and consequently only phase-lock to low-frequency components and/or the envelope of complex stimuli. This too suggests that a cochlear implant with its electrode array inserted into the basal turns may be better at coding low-frequency and envelope information in the periodicity of firings.

### 8.2.3 Stochastic auditory nerve response to electrical stimulation

An application of the analysis of stochastic auditory nerve response to electrical stimulation in Chapters 5 to 7 concerns both the effectiveness and the safety of high-rate pulsatile stimulation in cochlear implant stimulation strategies. Unless the inherent stochastic nature of the auditory nerve is taken into consideration, stimulus pulse-rates higher than an auditory fibre’s maximum discharge-rate appear to be without value. Moreover, it is possibly dangerous to the nerve—nerve damage studies have found that stimulation at high pulse-rates and high amplitudes does induce serious injury to the nerve (Tykocinski et al., 1995c; Tykocinski et al., 1995a). However, at lower

Table 8.1: Number of fibres with a discharge probability per pulse above 0.5 and number of fibres with a discharge probability above 0.9, from a model of 10 000 fibres, at intensities corresponding to different definitions of  $N_{\text{ucl}}$ .

Mode =		Bipolar			Monopolar		
$N_{\text{ucl}} =$		100	500	1 000	100	500	1 000
100 pps	pr > 0.5	1	32	77	0	0	6
	pr > 0.9	0	5	29	0	0	0
150 pps	pr > 0.5	0	18	47	0	0	0
	pr > 0.9	0	3	13	0	0	0

stimulus amplitudes which are within the animal's normal loudness range, nerve damage is not observed (Xu et al., 1997). One plausible explanation for these results is directly related to the importance of the stochastic nature of auditory nerve. If the nerve is stimulated with a high pulse-rate, low amplitude stimulus it would be expected that each fibre's discharge probability per pulse would be low, i.e., the fibre's discharge rate would be much lower than the stimulus's pulse rate (see Chapter 7). As a consequence, no damage to the fibres would occur. However, if the stimulus was high in amplitude, fibre discharge probabilities per pulse would become high, i.e., fibre discharge rates would indeed become so high (see Chapter 7) that they would cause injury to the nerve.

The stochastic model output has been analysed to determine what discharge probabilities per pulse are reached on individual fibres in what can be considered to be relatively extreme cases. Examples of discharge probability distributions are shown in Figure 8-1, for stimulation with a single electrode at intensities corresponding to 'uncomfortable loudness'. In Chapter 6 uncomfortable loudness was defined as a fixed number of neural discharges ( $N_{\text{ucl}}$ ) occurring within the temporal integration window of 100 ms. Table 8.1 shows the number of fibres with a discharge probability per pulse above 0.5 and the number of fibres with a discharge probability above 0.9, from a model of 10 000 fibres, for pulse trains of 100 and 150 pps at intensities corresponding to  $N_{\text{ucl}} = 100, 500$  and 1 000 spikes, for bipolar and monopolar electrode configurations. From the predictions of dynamic range and intensity difference limen presented in Section 6.3 it would appear that uncomfortable loudness ( $N_{\text{ucl}}$ ) is on the order of 100 to 1 000 spikes for the model. Normal listening levels are typically set somewhat below uncomfortable loudness levels, such that, even for continuous stimulation at high levels, speech processors are probably only producing discharges on the order of 100 to 500 spikes per 100 ms.

For bipolar stimulation, a 100 pps pulse-train of duration greater than 100 ms (i.e., 10 pulses occur within the temporal integration window) causes just 1 fibre out of 10 000 to have a discharge probability per pulse above 0.5 at  $N_{\text{ucl}} = 100$  spikes and just 32 fibres at  $N_{\text{ucl}} = 500$  spikes (out of

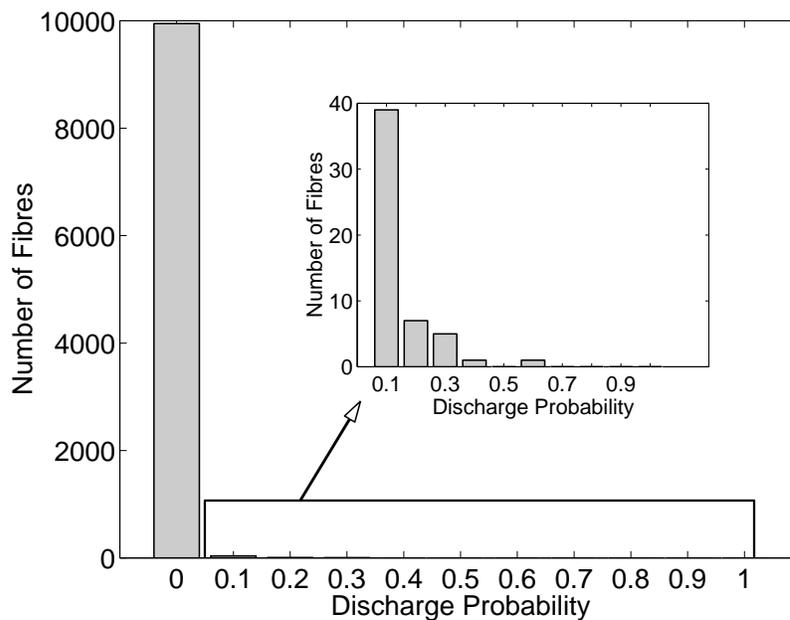
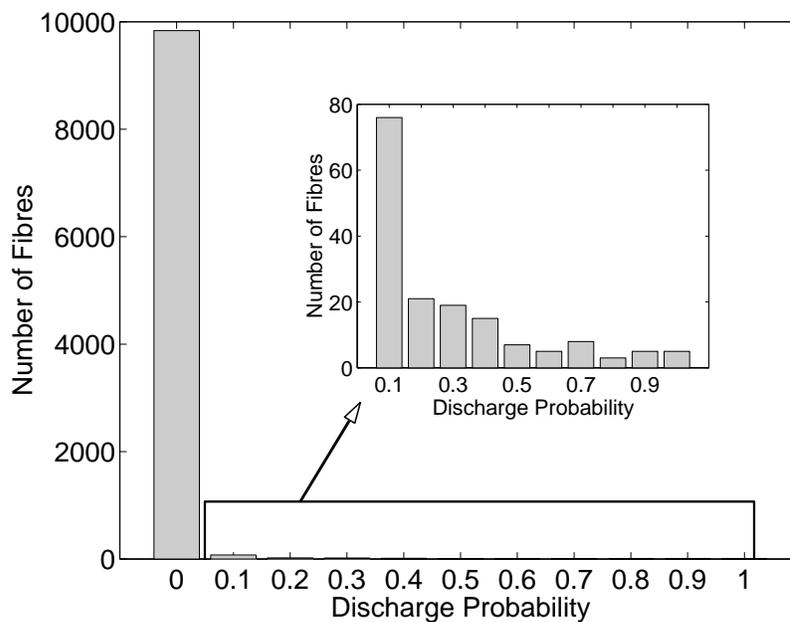
(a)  $N_{ucl} = 100$  spikes(b)  $N_{ucl} = 500$  spikes

Figure 8-1: Discharge probability histograms for a model of 10 000 fibres in response to a 100 pps pulse-train with a bipolar electrode configuration, at intensities corresponding to  $N_{ucl} = 100$  and 500 spikes. Insets zoom in on the region of the histogram for discharge probabilities greater than zero.

these 32 fibres only 5 have a discharge probability above 0.9). Increasing the pulse rate to 150 pps (15 pulses within the temporal integrator window) reduces the number of fibres above 0.5 discharge probability at  $N_{\text{ucl}} = 500$  spikes to 18 and out of these 18 fibres only 3 have a discharge probability above 0.9. For monopolar stimulation, *no* fibres have a discharge probability above 0.5 for a 100 pps pulse train, even at  $N_{\text{ucl}} = 500$  spikes.

From this analysis it appears that even in the most extreme cases individual spike probabilities are low except on a very few fibres. This means that the vast majority of fibres will not respond at all or only respond at a small fraction of the pulse rate. Furthermore, cochlear implants generally operate under far less extreme conditions. Firstly, most speech processing strategies stimulate using multiple electrodes within the 100 ms temporal integration window, such that the excitation contributing to the loudness of the stimulus is distributed between more fibres than when only a single electrode is used, i.e., discharge probabilities on individual fibres are lower in the case of multi-electrode stimulation. Secondly, pulse rates between 200 and 800 pps are typically being used in implants. Increasing the pulse rate causes more pulses to fall within the temporal integrator window such that uncomfortable loudness is reached at even lower individual discharge probabilities. Thirdly, these results are for a model of 10 000 fibres, which is a *conservative* estimate of the neural survival in implant users (Hinojosa and Marion, 1983). Greater neural survival also leads to a greater distribution of excitation amongst the fibres, again leading to reduced discharge probabilities on individual fibres. Finally, speech signals vary in their spectrum and envelope amplitude over time, such that continuous stimulation of the one population of AN fibres is unlikely to occur for extended periods of time. These results suggest that continuous, multi-electrode, pulsatile stimulation at moderate rates does *not* drive AN fibres beyond their normal physiological limits.

Several cochlear implant research groups (e.g., Rubinstein, Matsuoka, Miller and Abbas, 1997) have suggested that increasing the relative noise level of nerve fibres may improve the ‘temporal representation’ of the stimulus. It is argued that: (i) noisier fibres will reduce the occurrence of highly-synchronous discharge patterns which are not seen in responses to acoustical stimulation and (ii) noisier fibres will reduce refractory effects that distort the temporal representation of the stimulus. The results obtained in Chapters 5, 6 and 7, suggest several methods for increasing the relative noise level of fibres. This appears to be a very useful strategy, but may be entirely unnecessary. If an implant user’s nerve fibres discharge only at low probabilities during multi-channel continuous stimulation, refractory effects will be negligible because discharge rates will be low. Also, abnormally high synchronisation of discharges across fibres will not occur. For example, with a 1 000 pps stimulus and fibre discharge probabilities less than 0.1, relatively little temporal distortion due to refractory effects will occur, because the average interval between discharges within any fibre will be greater than 10 ms.

## 8.2.4 Future directions

The results of this thesis indicate the utility of mathematical and computational models in investigating the spatiotemporal coding of sound in the auditory nerve. In some cases the results suggest what hypothetical neural structures may be useful in processing and decoding certain sound features, prompting anatomical and physiological investigation of existing structures within the auditory pathways. In particular, the innervation of the cochlear nucleus by auditory nerve fibres needs further experimental investigation. Experimental data which reveal more about neural connectivity between auditory nerve fibres and cochlear nucleus neurons may help to validate the models or suggest improvements to them. Comparison of responses to acoustical and electrical stimulation in the same experimental animal may be particularly useful in this task, because the two modes of stimulation provide quite different patterns of spatiotemporal response.

Some of the results of this thesis have been tested by direct comparison to psychophysical data, while others have only been subjected to theoretical considerations. Development of psychophysical models such as the one derived in Chapter 6 for perception of loudness with a cochlear implant would permit direct prediction of further psychophysical data. For instance, the models of temporal frequency coding developed in Chapters 2 to 4, and other models proposed by theories of frequency coding, could be extended and combined with the model of auditory nerve response to electrical pulse trains derived in Chapter 7 to predict the effects of pulse rate on the pitch of electrical stimuli.

The approach taken in this thesis in comparing model predictions with psychophysical data has been to avoid “curve fitting”, i.e., finding model parameters which produce the best predictions. This has been done in order to demonstrate that parameter values inferred directly from physiological data may be sufficient to produce reasonable qualitative predictions of psychophysical data from many subjects. However, because of the reasonable qualitative accuracy of the model predictions, the next step in these comparisons would be to conduct more quantitative evaluations. These could involve fitting the models to psychophysical data of individual subjects, in order to investigate how model parameters and the corresponding physiology may explain inter-subject variability. Furthermore, the psychophysical models utilised have been for relatively simple central processing and an ideal or efficient observer—more complex psychophysical models of detection, discrimination and estimation by human subjects may improve the model predictions.

Extensions and improvements could certainly be made to the modelling methods, and have been discussed in some of the individual chapters. Great scope exists particularly for mathematical and computational modelling of stochastic and nonlinear properties of auditory nerve response to electrical stimulation. However, their usefulness may be limited by the conceptual and computational complexity introduced and by what experimental data is available to validate them. In particular, the results of the investigation of stochastic activity in auditory nerve response to

electrical stimulation suggest that much of the available physiological data may have been collected for stimulus intensities which are higher than those typically produced by cochlear implant speech processing strategies. Physiological data collected at appropriate intensities may assist further in understanding the relationship between perception of stimuli generated by cochlear implants and the underlying spatiotemporal patterns of auditory nerve response.

# Appendix A

## Proofs for Chapter 3

### A.1 Proof of Claim 3.2.1

Let the absolute difference between  $\mu_I(t)$  and  $\mu_S(t)$  equal  $\theta$ .

$$\left| \frac{N}{d} \int_0^d \lambda(x, t) dx - \sum_{n=0}^{N-1} \lambda_n(t) \right| = \theta \quad (\text{A.1})$$

Multiplying both sides of (A.1) by  $\frac{d}{N}$  gives:

$$\left| \int_0^d \lambda(x, t) dx - \sum_{n=0}^{N-1} \frac{d}{N} \lambda_n(t) \right| = \frac{d}{N} \theta \quad (\text{A.2})$$

The left hand side of (A.2) is the difference between the integral over distance  $d$  of a function continuous in space and the sum of  $N$  rectangles each of width  $\frac{d}{N}$  and of height  $\lambda_n(t)$ . The greatest difference will occur where the absolute slope of the rate  $\lambda(x, t)$  is at a maximum. The absolute slope of the rate is bound by  $\left| \frac{\delta \lambda(x, t)}{\delta x} \right| \leq |\alpha B|$ , and summing the maximal errors of the individual rectangles then dividing by the mean gives (3.15).

□

### A.2 Proof of Lemma 3.2.2

Equations (3.8) and (3.9) arise from Vannucci and Teich (1978).

#### Proof of (3.7)

From Cox (1962):

$$\text{E}[N_n(t, \tau)] = \frac{s_0}{1 + as_0} \tau \quad (\text{A.3})$$

and

$$\text{var}[N_n(t, \tau)] = \Psi_n(t, \tau) - \text{E}[N_n(t, \tau)] - \text{E}^2[N_n(t, \tau)] \quad (\text{A.4})$$

where

$$\begin{aligned} \Psi_n(t, \tau) &= \mathcal{L}^{-1} \{ \Psi_n^*(t, s); s \rightarrow \tau \} \\ &= \mathcal{L}^{-1} \left\{ \frac{2s_0}{s^2(1+as_0) \left[ 1 - \frac{s_0 e^{-as}}{s+s_0} \right]}; s \rightarrow \tau \right\} \end{aligned}$$

Thus:

$$\begin{aligned} \Psi_n^*(t, s) &= \frac{2s_0}{1+as_0} \frac{1}{s^2} \frac{s+s_0}{s+s_0} \frac{1}{(1-e^{-as})} \\ &= \frac{2s_0}{1+as_0} \frac{1}{s^2} \sum_{k=0}^{\infty} \left( \frac{s_0 e^{-as}}{s+s_0} \right)^k \\ &= \frac{2s_0}{1+as_0} \sum_{k=0}^{\infty} (s_0 e^{-as})^k \frac{1}{s^2} \frac{1}{(s+s_0)^k} \end{aligned} \quad (\text{A.5})$$

Taking the inverse Laplace transform of (A.5) gives:

$$\begin{aligned} \Psi_n(t, \tau) &= \frac{2s_0}{1+as_0} \sum_{k=1}^{\infty} \frac{s_0^k}{(k-1)!} U(\tau - aj) \int_0^{\tau-ak} [\tau - ak - \gamma] \gamma^{k-1} e^{-s_0 \gamma} d\gamma \\ &\quad + \frac{2s_0}{1+as_0} \tau \end{aligned} \quad (\text{A.6})$$

Substituting for (A.3) and (A.6) in (A.4) gives (3.7).

□

### A.3 Proof of Lemma 3.3.1

Follows immediately from the Tchebycheff Inequality—see (5-57) in Papoulis (1991) for instance.

# Appendix B

## Proofs for Chapter 5

### B.1 Proof of Lemma 5.3.1

The probability of discharge during pulse  $n$  is equal to the probability that  $V_{\text{stim}}(n)$  is greater than the noise modified threshold during pulse  $n$ .

$$\begin{aligned} p(n) &= \Pr \{V_{\text{stim}}(n) \geq V_{\text{thr}}(n) + V_{\text{noise}}(n)\} \\ &= \Pr \{V_{\text{noise}}(n) \leq V_{\text{stim}}(n) - V_{\text{thr}}(n)\} \\ &= \int_{-\infty}^{V_{\text{stim}}(n) - V_{\text{thr}}(n)} f(x) dx \end{aligned}$$

where  $f(x)$  is the probability distribution of  $V_{\text{noise}}(n)$ .

For  $V_{\text{noise}}(n)$  distributed normally with a mean of zero and standard deviation  $\sigma$  ( $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2}$ ):

$$\begin{aligned} p(n) &= \int_{-\infty}^{V_{\text{stim}}(n) - V_{\text{thr}}(n)} \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} dx \\ &= \frac{1}{2} \left( 1 + \operatorname{erf} \left( \frac{V_{\text{stim}}(n) - V_{\text{thr}}(n)}{\sqrt{2}\sigma} \right) \right) \end{aligned}$$

where erf is the integrated-Gaussian, or Error Function:

$$\operatorname{erf}(x) \triangleq \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$$

□

# Appendix C

## Proofs for Chapter 6

### C.1 Proof of Lemma 6.2.1

Consider the detection/discrimination system described in Lemma 6.2.1. The probability of choosing correctly is equal to the probability that the neural response to the stimulus with the greater intensity ( $X_2$ ) is greater than that to the lower intensity stimulus ( $X_1$ ), plus the probability that a correct guess is made if the responses are identical. That is:

$$\Pr\{\text{Choosing correctly}\} = \Pr\{X_2 > X_1\} + \frac{1}{2} \Pr\{X_2 = X_1\}$$

For the random variables  $X_1$  and  $X_2$  with integer values bound by  $[0, X_{\max}]$  and probability mass functions  $f_1$  and  $f_2$ :

$$\Pr\{X_2 > X_1\} = \sum_{n=0}^{X_{\max}} f_1(n) \sum_{m=n+1}^{X_{\max}} f_2(m)$$

and

$$\Pr\{X_1 = X_2\} = \sum_{n=0}^{X_{\max}} f_1(n)f_2(n)$$

□

# Appendix D

## Proofs for Chapter 7

### D.1 Proof of Lemma 7.3.1

An expression for  $g(n_i|j)$  follows directly from (5.2), for *Case 1* where  $V_{\text{noise}}$  is not affected by the refractory function, and *Case 2* where  $V_{\text{noise}}$  increases during the refractory period proportional to the refractory modified threshold potential divided by the resting threshold potential ( $m(n_i|j)$ ).

*Case 1:*

Substituting  $V_{\text{thr}}(n) + V_{\text{refr}}(n_i|j)$  for  $V_{\text{thr}}(n)$  in (5.2) gives (7.4).

*Case 2:*

$$\begin{aligned} g(n_i|j) &= \Pr\{V_{\text{stim}}(n) \geq V_{\text{thr}}(n) + V_{\text{refr}}(n_i|j) + V_{\text{noise}}(n_i|j)\} \\ &= \Pr\{V_{\text{stim}}(n) \geq V_{\text{thr}}(n) + V_{\text{refr}}(n_i|j) + V_{\text{noise}}(n) \times m(n_i|j)\} \\ &= \Pr\{V_{\text{noise}}(n) \leq V_{\text{stim}}(n)/m(n_i|j) - V_{\text{thr}}(n)\} \\ &= \int_{-\infty}^{V_{\text{stim}}(n)/m(n_i|j) - V_{\text{thr}}(n)} f(x) dx \\ &= \frac{1}{2} \left( 1 + \operatorname{erf} \left( \frac{V_{\text{stim}}(n)/m(n_i|j) - V_{\text{thr}}(n)}{\sqrt{2}\sigma} \right) \right) \end{aligned}$$

For a monotonically decreasing refractory (and, in *Case 2*, noise) modified threshold, a discharge will occur during pulse  $n$  if  $V_{\text{stim}}(n)$  is greater than the refractory (and, in *Case 2*, noise) modified threshold by the final bin of pulse  $n$ . Therefore  $p(n|j)$  is equal to  $g(n_i|j)$  when  $i$  is the final bin of pulse  $n$ , giving (7.6).

□

## D.2 Proof of Lemma 7.3.2

If  $f(n|j)$  is the probability that the *first subsequent* discharge occurs during pulse  $n$ , given that the last discharge occurred  $n$  pulses ago in bin  $j$ :

For  $n = 0$ :

$$\begin{aligned} f(n|j) &= \Pr\{\text{next discharge occurs during same pulse} \\ &\quad | \text{last discharge occurred in bin } j\} \\ &= p(0|j) \end{aligned}$$

For  $n > 0$ :

$$\begin{aligned} f(n|j) &= \Pr\{\text{no discharge before } n \\ &\quad | \text{last discharge occurred } n \text{ pulses ago in bin } j\} \\ &\quad \times \Pr\{\text{discharge during pulse } n \\ &\quad | \text{last discharge occurred } n \text{ pulses ago in bin } j\} \\ &= Q(n-1|j)p(n|j) \end{aligned}$$

□

## D.3 Proof of Lemma 7.3.3

$$E[r|j] = \sum_{k=0}^{\infty} kf(k|j)$$

Consider the case where the pulse train is of infinite duration and there exists some time  $L$  such that refractory effects are negligible for all times  $k \geq L$  such that  $p(k|j) = p, \forall k \geq L$ . Then:

$$E[r|j] = \sum_{k=0}^{L-1} kf(k|j) + \sum_{k=L}^{\infty} kf(k|j)$$

For  $k \geq L$ ,

$$\begin{aligned} f(k|j) &= \Pr\{\text{no discharge before } L\} \\ &\quad \times \Pr\{\text{no discharge between } L \text{ and } k\} \\ &\quad \times \Pr\{\text{discharge during } k\} \\ &= Q(L-1|j)(1-p)^{k-L}p \end{aligned}$$

Then:

$$\begin{aligned} \mathbb{E}[r|j] &= \sum_{k=0}^{L-1} k f(k|j) + \sum_{k=L}^{\infty} k Q(L-1|j)(1-p)^{k-L} p \\ &= \sum_{k=0}^{L-1} k f(k|j) + Q(L-1|j) \left( L-1 + \frac{1}{p} \right) \end{aligned}$$

Likewise:

$$\begin{aligned} \text{var}[r|j] &= \mathbb{E}[r^2|j] - \mathbb{E}[r|j]^2 \\ &= \sum_{k=0}^{\infty} k^2 f(k|j) - \mathbb{E}[r|j]^2 \\ &= \sum_{k=0}^{L-1} k^2 f(k|j) + \sum_{k=L}^{\infty} k^2 f(k|j) - \mathbb{E}[r|j]^2 \\ &= \sum_{k=0}^{L-1} k^2 f(k|j) \\ &\quad + \sum_{k=L}^{\infty} k^2 Q(L-1|j)(1-p)^{k-L} p - \mathbb{E}[r|j]^2 \\ &= \sum_{k=0}^{L-1} k^2 f(k|j) \\ &\quad + Q(L-1|j) \left( \left( L-1 + \frac{1}{p} \right)^2 + \frac{1-p}{p^2} \right) - \mathbb{E}[r|j]^2 \end{aligned}$$

□

## D.4 Proof of Lemma 7.3.4

$$\begin{aligned} B_{ij} &= \Pr\{\text{next discharge occurs in any bin } i \\ &\quad | \text{last discharge occurred in any bin } j\} \\ &= \sum_{k=0}^{\infty} \Pr\{\text{no discharge before } k | j\} \\ &\quad \times \Pr\{\text{discharge occurs in bin } i \text{ of pulse } k | j\} \\ &= \sum_{k=0}^{\infty} Q(k-1|j) p(k_i|j) \end{aligned}$$

Consider the case where the pulse train is of infinite duration and there exists some time  $L$  such that refractory effects are negligible for all times  $k \geq L$  such that  $p(k|j) = p, \forall k \geq L$ . Then:

$$B_{ij} = \sum_{k=0}^{L-1} Q(k-1|j)p(k_i|j) + \sum_{k=L}^{\infty} Q(k-1|j)p(k_i|j)$$

For  $k \geq L$ ,

$V_{\text{refr}}(k_i)$  is zero for all values of  $i$ , and therefore if a discharge occurs it will always occur in the first bin of pulse  $k$ . Thus, the probability that a discharge occurs in the first bin is equal to the probability of a discharge occurring in that pulse, i.e.,  $p(k_1|j) = p(k|j)$ .

It follows that:

$$\begin{aligned} \sum_{k=L}^{\infty} Q(k-1|j)p(k_i|j) &= \sum_{k=L}^{\infty} Q(k-1|j)p(k|j) \\ &= \sum_{k=L}^{\infty} Q(k-1|j) \\ &\quad - \sum_{k=L}^{\infty} Q(k-1|j) \times (1-p(k|j)) \\ &= \sum_{k=L}^{\infty} Q(k-1|j) - \sum_{k=L}^{\infty} Q(k|j) \\ &= \sum_{k=L}^{\infty} Q(k-1|j) - \sum_{k=L+1}^{\infty} Q(k-1|j) \\ &= Q(L-1|j) \end{aligned}$$

For all  $i$ :

$$\sum_{k=L}^{\infty} Q(k-1|j)p(k_i|j) = \begin{cases} Q(L-1|j) & \text{for } i = 1 \\ 0 & \text{for } i \neq 1 \end{cases}$$

Thus:

$$B_{ij} = \sum_{k=0}^{L-1} Q(k-1|j)p(k_i|j) + Q(L-1|j)\delta_{i=1}$$

With a monotonically decreasing refractory and noise modified threshold, the probability of discharge in bin  $i$  of pulse  $n$ , given that the last discharge occurred  $n$  pulses previously in bin  $j$ , is equal to the probability that  $V_{\text{stim}}(n)$  is greater than the refractory and noise modified threshold in bin  $i$  of pulse  $n$ , but was less than the modified threshold in the previous bin,  $i-1$ .

For *Case 1*:

$$\begin{aligned} p(n_i|j) &= \Pr\{V_{\text{thr}}(n) + V_{\text{refr}}(n_i|j) + V_{\text{noise}}(n) \leq V_{\text{stim}}(n) \\ &\quad \leq V_{\text{thr}}(n) + V_{\text{refr}}(n_{i-1}|j) + V_{\text{noise}}(n)\} \\ &= \Pr\{V_{\text{stim}}(n) - V_{\text{thr}}(n) - V_{\text{refr}}(n_{i-1}|j) \leq V_{\text{noise}}(n) \end{aligned}$$

$$\begin{aligned}
& \leq V_{\text{stim}}(n) - V_{\text{thr}}(n) - V_{\text{refr}}(n_i|j) \} \\
& = \int_{V_{\text{stim}}(n) - V_{\text{thr}}(n) - V_{\text{refr}}(n_{i-1}|j)}^{V_{\text{stim}}(n) - V_{\text{thr}}(n) - V_{\text{refr}}(n_i|j)} f(x) dx \\
& = \int_{-\infty}^{V_{\text{stim}}(n) - V_{\text{thr}}(n) - V_{\text{refr}}(n_i|j)} f(x) dx \\
& \quad - \int_{-\infty}^{V_{\text{stim}}(n) - V_{\text{thr}}(n) - V_{\text{refr}}(n_{i-1}|j)} f(x) dx \\
& = g(n_i|j) - g(n_{i-1}|j)
\end{aligned}$$

For *Case 2*:

$$\begin{aligned}
p(n_i|j) &= \Pr\{V_{\text{thr}}(n) + V_{\text{refr}}(n_i|j) + V_{\text{noise}}(n_i|j) \leq V_{\text{stim}}(n) \\
& \leq V_{\text{thr}}(n) + V_{\text{refr}}(n_{i-1}|j) + V_{\text{noise}}(n_{i-1}|j) \} \\
& = \Pr\{V_{\text{thr}}(n) + V_{\text{refr}}(n_i|j) + V_{\text{noise}}(n) \times m(n_i|j) \leq V_{\text{stim}}(n) \\
& \leq V_{\text{thr}}(n) + V_{\text{refr}}(n_{i-1}|j) + V_{\text{noise}}(n) \times m(n_i|j) \} \\
& = \Pr\{V_{\text{stim}}(n)/m(n_{i-1}|j) - V_{\text{thr}}(n) \leq V_{\text{noise}}(n) \\
& \leq V_{\text{stim}}(n)/m(n_i|j) - V_{\text{thr}}(n) \}
\end{aligned}$$

Following Assumption 7.3.1,  $V_{\text{refr}}(n_{i-1}|j) > V_{\text{refr}}(n_i|j) \geq 0$ . Therefore:

$$\frac{1}{m(n_i|j)} > \frac{1}{m(n_{i-1}|j)}$$

and

$$\begin{aligned}
p(n_i|j) &= \int_{-\infty}^{V_{\text{stim}}(n)/m(n_i|j) - V_{\text{thr}}(n)} f(x) dx \\
& \quad - \int_{-\infty}^{V_{\text{stim}}(n)/m(n_{i-1}|j) - V_{\text{thr}}(n)} f(x) dx \\
& = g(n_i|j) - g(n_{i-1}|j)
\end{aligned}$$

Thus, for both *Case 1* and *Case 2*:

$$p(n_i|j) = g(n_i|j) - g(n_{i-1}|j)$$

where  $g(n_i|j)$  is described by (7.4) for *Case 1* and by (7.5) for *Case 2*.

□

## D.5 Proof of Lemma 7.3.5

Given that the probability of discharge is non-zero for all pulses and that the pulse train is of infinite duration, a discharge must occur at some stage in the pulse train. Consequently, the probability of the next discharge occurring in any bin  $i$  summed over all bins must equal one, i.e., the sum of all elements in each column of  $\mathbf{B}$  equals one.

The sum of all elements in each row of  $\mathbf{B}^T$  is then equal to one. Therefore, defining  $\vec{\mathbf{1}}$  as a column vector whose elements are all equal to one:

$$\mathbf{B}^T \vec{\mathbf{1}} = \vec{\mathbf{1}}$$

$\mathbf{B}^T$  must therefore have an eigenvalue which is equal to one.

The eigenvalues of  $\mathbf{B}$  are equal to the eigenvalues of  $\mathbf{B}^T$ , and therefore  $\mathbf{B}$  must also have an eigenvalue which is equal to one.

□

# Bibliography

- Anderson, D. J. (1973). Quantitative model for the effects of stimulus frequency upon synchronization of auditory nerve discharges, *J. Acoust. Soc. Am.* **54**: 361–4.
- Battmer, R.-D., Gnadeberg, D., von Wallenberg, E., Lehnhardt, E. and Allum, D. J. (1993). A study of monopolar and bipolar stimulation modes with a modified Nucleus Mini-22 cochlear implant, *Adv. Otorhinolaryngol.* **48**: 9–16.
- Bi, Q. (1989). A closed-form solution for removing the dead time effects from the poststimulus time histograms, *J. Acoust. Soc. Am.* **85**(6): 2504–13.
- Bostock, H. (1983). The strength-duration relationship for excitation of myelinated nerve: Computed dependence on membrane parameters, *J. Physiol.* **341**: 59–74.
- Bruce, I. C., Irlicht, L. S., White, M. W., O’Leary, S. J., Dynes, S., Javel, E. and Clark, G. M. (1997a). A stochastic model of the electrically stimulated auditory nerve: Single-pulse response, *Submitted to IEEE Trans. Biomed. Eng.*
- Bruce, I. C., Irlicht, L. S., White, M. W., O’Leary, S. J., Dynes, S., Javel, E. and Clark, G. M. (1997b). A stochastic model of the electrically stimulated auditory nerve: Pulse-train response, *Submitted to IEEE Trans. Biomed. Eng.*
- Bruce, I. C., White, M. W., Irlicht, L. S., O’Leary, S. J. and Clark, G. M. (1997). The effects of stochastic neural activity in a model predicting loudness perception with cochlear implants: Low-rate stimulation, *Submitted to IEEE Trans. Biomed. Eng.*
- Burkitt, A. N. and Clark, G. M. (1997). Analysis of integrate and fire neurons: Synchronization of synaptic input and spike output, *Submitted to Neural Computation* .
- Butikofer, R. and Lawrence, P. D. (1979). Electrocutaneous nerve stimulation. II: Stimulus waveform selection, *IEEE Trans. Biomed. Eng.* **BME-26**: 69–75.
- Carney, L. H. (1992). Modelling the sensitivity of cells in the anteroventral cochlear nucleus to spatiotemporal discharge patterns, *Phil. Trans. R. Soc. Lond.* **B336**: 403–406.

- Carney, L. H. (1994). Spatiotemporal encoding of sound level: Models for normal encoding and recruitment of loudness, *Hear. Res.* **76**: 31–44.
- Clark, G. M. (1987). *The University of Melbourne–Nucleus Multi-Electrode Cochlear Implant*, Karger, Basel.
- Clark, G. M. (1996). Electrical stimulation of the auditory nerve: The coding of frequency, the perception of pitch and the development of cochlear implant speech processing strategies for profoundly deaf people, *Clinical and Experimental Pharmacology and Physiology* **23**: 766–776.
- Clark, G. M. (1997a). Advances in cochlear implant speech processing, in G. M. Clark (ed.), *Cochlear Implants. XVI World Congress of Otorhinolaryngology Head and Neck Surgery*, Monduzzi Editore, Bologna, pp. 9–15.
- Clark, G. M. (1997b). Cochlear implant research directions, in G. M. Clark (ed.), *Cochlear Implants. XVI World Congress of Otorhinolaryngology Head and Neck Surgery*, Monduzzi Editore, Bologna, pp. 55–60.
- Clark, G. M., Carter, T. D., Maffi, C. L. and Shepherd, R. K. (1995). Temporal coding of frequency: Neuron firing probabilities for acoustic and electrical stimulation of the auditory nerve, *Ann. Otol. Rhinol. Laryngol.* **104**(Suppl. 166): 109–111.
- Clark, G. M., Dowell, R. C., Cowan, R. S., Pyman, B. C. and Webb, R. L. (1995). Multicentre evaluations of speech perception in adults and children with the Nucleus (Cochlear) 22-channel cochlear implant, in M. Portmann (ed.), *Transplants and Implants in Otology III. Proceedings of the 3rd International Symposium of Transplants and Implants in Otology, Bordeaux, France*, Kugler Publications, The Netherlands, pp. 1–11.
- Clark, G. M., Irlicht, L. S. and Carter, T. D. (1996). A neural model for the time-period coding of frequency for acoustic and electric stimulation, *Presented at the 16th Annual Meeting of the Australian Neuroscience Society*.
- Clay, J. R. and DeFelice, L. J. (1983). Relationship between membrane excitability and single channel open-close kinetics, *Biophysical Journal* **42**: 151–157.
- Cohen, L. T., Saunders, E., Busby, P. A. and Clark, G. M. (1997). Place information in cochlear implants: An integrated approach employing psychophysics and modelling, *Program and Abstracts of 1997 Conference on Implantable Auditory Prostheses. Asilomar Conference Center, Pacific Grove, California*.
- Collins, J. J., Chow, C. C. and Imhoff, T. T. (1995). Stochastic resonance without tuning, *Nature* pp. 236–237.

- Colombo, J. and Parkins, C. W. (1987). A model of electrical excitation of the mammalian auditory-nerve neuron, *Hear. Res.* **31**: 287–312.
- Cox, D. R. (1962). *Renewal Theory*, Science Paperbacks, Chapman and Hall, New York. ISBN 0-412-20570-X.
- Dallos, P., Popper, A. N. and Fay, R. R. (1996). *The Cochlea*, Springer, New York.
- Dean, D. and Lawrence, P. D. (1983). Application of phase analysis of Frankenhaeuser–Huxley equations to determine threshold stimulus amplitudes, *IEEE Trans. Biomed. Eng.* **BME-30**(12): 810–18.
- Delgutte, B. (1996). Physiological models for basic auditory percepts, in H. L. Hawkins and T. McMullen (eds), *Auditory Computation*, Springer Verlag, New York, chapter 5.
- Derksen, H. E. and Verveen, A. A. (1966). Fluctuations of resting neural membrane potential, *Science* **151**: 1388–1389.
- Devore, J. L. (1987). *Probability and Statistics for Engineering and the Sciences*, second edn, Brooks/Cole Publishing Company, Monterey, California.
- Dodge, F. A. (1961). Ionic permeability changes underlying nerve excitation, in A. M. Shanes (ed.), *Biophysics of Physiological and Pharmacological Actions*, Publication no. 69 of the American Association for the Advancement of Science, Washington D.C., p. 119.
- Dynes, S. (1996). *Discharge Characteristics of Auditory Nerve Fibers for Pulsatile Electrical Stimuli*, PhD thesis, Massachusetts Institute of Technology.
- Eddington, D. K. (1980). Speech discrimination in deaf subjects with cochlear implants, *J. Acoust. Soc. Am.* **68**: 885–891.
- Eddington, D. K., Dobbelle, W. H., Brackmann, D. E., Mladejovsky, M. G. and Parkin, J. L. (1978). Auditory prostheses research with multiple channel intracochlear stimulation in man, *Ann. Otol. Rhinol. Laryngol.* **87**(6): 1–39.
- Edwards, B. W. and Wakefield, G. H. (1990). On the statistics of binned neural point processes: The Bernoulli approximation and AR representation of the PST histogram, *Biol. Cybern.* **64**: 145–53.
- Eggermont, J. J. (1990). On the pathophysiology of tinnitus: A review and a peripheral model, *Hear. Res.* **48**(1–2): 111–123.
- Evans, E. F. (1981). The dynamic range problem: Place and timing coding at the level of the cochlear nerve and nucleus, in J. Syka and L. Aitkin (eds), *Neuronal Mechanisms of Hearing*, Plenum, New York, pp. 69–95.

- Feller, W. (1968). *An Introduction to Probability Theory and Its Applications*, Vol. 1, third edn, Wiley Series in Probability and Mathematical Statistics, New York.
- Ferguson, W. D., Smith, D. W., Finley, C. C., Pfingst, B. E. and Collins, L. M. (1998). Prediction of the variance in behavioral thresholds using a stochastic model of electrical stimulation, *ARO Midwinter Meeting Abstracts*.
- Finley, C. C., Wilson, B. S. and White, M. W. (1990). Models of neural responsiveness to electrical stimulation, in J. M. Miller and F. A. Spelman (eds), *Cochlear Implants: Models of the Electrically Stimulated Ear*, Springer-Verlag, New York, pp. 55–96.
- Frankenhaeuser, B. (1957). A method for recording resting and action potentials in the isolated myelinated nerve fibre of the frog, *J. Physiol. (Lond.)* **135**: 550.
- Frankenhaeuser, B. and Huxley, A. (1964). The action potential in the myelinated nerve fiber of *xenopus laevis* as computed on the basis of voltage clamp data, *J. Physiol.* **171**: 302–15.
- Frijns, J. H. M. (1995). *Cochlear Implants - A Modelling Approach*, PhD thesis, Leiden University, The Netherlands.
- Gantz, B. J., McCabe, F., Tyler, R. S. and Preece, J. P. (1987). Evaluation of four cochlear implant designs, *Ann. Otol. Rhinol. Laryngol.* **96**(Suppl. 128): 145–147.
- Gaumond, R. P., Molnar, C. E. and Kim, D. O. (1982). Stimulus and recovery dependence of cat cochlear nerve fiber spike discharge probability, *J. Neurophysiol.* **48**: 856–73.
- Gleich, O. and Wilson, S. (1993). The diameter of guinea pig auditory nerve fibres: Distribution and correlation with spontaneous rate, *Hear. Res.* **71**: 69–79.
- Goldstein, J. L. and Sruлович, P. (1977). Auditory-nerve spike intervals as an adequate basis for aural spectrum analysis, in E. F. Evans and J. P. Wilson (eds), *Psychophysics and Physiology of Hearing*, Academic, London, pp. 337–345.
- Green, D. M. and Swets, J. A. (1973). *Signal Detection Theory and Psychophysics*, John Wiley and Sons 1969, reprinted by R. E. Krieger Publ. Co., New York.
- Gregson, R. A. M. (1988). *Nonlinear Psychophysical Dynamics*, Lawrence Erlbaum Associates, Inc., Hillsdale, New Jersey.
- Guiaşu, S. (1977). *Information Theory with Applications*, McGraw-Hill Inc., London.
- Hellman, W. S. and Hellman, R. P. (1990). Intensity discrimination as the driving force for loudness. application to pure tones in quiet, *J. Acoust. Soc. Am.* **87**: 1255–65.
- Hill, A. V. (1936). Excitation and accommodation in nerve, *Proc. R. Soc. B* **119**: 305–355.

- Hinojosa, R. and Marion, M. (1983). Histopathology of profound sensorineural deafness, *Ann. N.Y. Acad. Sci.* **405**: 459–484.
- Hochmair-Desoyer, I. J., Hochmair, E. S., Fischer, R. E. and Burian, K. (1980). Cochlear prostheses in use: Recent speech comprehension results, *Arch. Otorhinolaryngol.* **229**: 81–98.
- Hochmair-Desoyer, I. J., Hochmair, E. S., Motz, H. and Rattay, F. (1984). A model for the electrostimulation of the nervus acusticus, *Neuroscience* **13**(2): 553–562.
- Hocking, J., Brown, M., and Clark, G. M. (1997). Temporal coding in auditory neurons to electrical stimulation, in G. M. Clark (ed.), *Cochlear Implants. XVI World Congress of Otorhinolaryngology Head and Neck Surgery*, Monduzzi Editore, Bologna, pp. 115–118.
- Hodgkin, A. and Huxley, A. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve, *J. Physiol.* **117**: 500–44.
- Hodgkin, A. L. (1961). *The Conduction of the Nervous Impulse*, Liverpool University Press.
- Hoopen, M. T. and Verveen, A. A. (1963). Nerve-model experiments on fluctuation in excitability, *Prog. Brain Research* **2**: 8–21.
- Horst, J. W., Javel, E. and Farley, G. R. (1986). Coding of spectral fine structure in the auditory nerve. I. Fourier analysis of period and interspike interval histograms., *J. Acoust. Soc. Am.* **79**(2): 398–416.
- Horst, J. W., Javel, E. and Farley, G. R. (1990). Coding of spectral fine structure in the auditory nerve. II: Level-dependent nonlinear responses, *J. Acoust. Soc. Am.* **88**(6): 2656–2681.
- House, W. F. and Berliner, K. I. (1982). Cochlear implants: Progress and perspectives., *Ann. Otol. Rhinol. Laryngol.* **91**(Suppl. 91).
- Irlicht, L. S. and Clark, G. M. (1996). Control strategies for neurons modeled by self-exciting point processes, *J. Acoust. Soc. Am.* **100**(5): 3237–3247.
- Javel, E. and Mott, J. B. (1988). Physiological and psychophysical correlates of temporal processes in hearing, *Hear. Res.* **34**: 275–294.
- Javel, E., Tong, Y. C., Shepherd, R. K. and Clark, G. M. (1987). Responses of cat auditory nerve fibers to biphasic electrical current pulses, *Ann. Otol. Rhinol. Laryngol.* **96**(Suppl. 128): 26–30.
- Jenison, R. L., Greenberg, S., Kluender, K. R. and Rhode, W. S. (1991). A composite model of the auditory periphery for the processing of speech based on the filter response functions of single auditory-nerve fibers, *J. Acoust. Soc. Am.* pp. 773–786.

- Johnson, D. (1980). The relationship between spike rate and synchrony in responses to auditory-nerve fibers to single tones., *J. Acoust. Soc. Am.* **68**: 1115–22.
- Johnson, D. H. and Kiang, N. Y. (1976). Analysis of discharges recorded simultaneously from pairs of auditory nerve fibers, *Biophys. J.* **16**(7): 719–734.
- Johnson, D. and Swami, A. (1983). The transmission of signals by auditory-nerve fiber discharge patterns, *J. Acoust. Soc. Am.* **74**(2): 493–501.
- Johnson, N. L., Kotz, S. and Kemp, A. W. (1992). *Univariate Discrete Distributions*, Wiley Series in Probability and Mathematical Statistics, second edn, John Wiley & Sons, Inc., New York.
- Jones, K. and Tubis, A. (1985). On the extraction of the signal-excitation function from a non-Poisson cochlear neural spike train, *J. Acoust. Soc. Am.* **78**(1): 90–4.
- Joris, P. X., Carney, L. H., Smith, P. H. and Yin, T. C. T. (1994). Enhancement of neural synchronization in the anteroventral cochlear nucleus. I. Responses to tones at the characteristic frequency, *J. Neurophysiol.* **71**: 1022–1036.
- Joris, P. X., Smith, P. H. and Yin, T. C. T. (1994). Enhancement of neural synchronization in the anteroventral cochlear nucleus. II. Responses in the tuning curve tail, *J. Neurophysiol.* **71**: 1037–1051.
- Kay, S. M. (1993). *Fundamentals of Statistical Signal Processing - Estimation Theory*, Prentice Hall Signal Processing Series, PTR Prentice Hall, New Jersey.
- Keeping, E. S. (1962). *Introduction to Statistical Inference*, D. Van Nostrand Company, New York.
- Kiang, N. Y. S., Moxon, E. C. and Levine, R. A. (1970). Auditory-Nerve activity in cats with normal and abnormal cochleas, in G. E. W. Wolstenholme and J. Knight (eds), *Sensorineural Hearing Loss*, A Ciba Foundation Symposium, J. & A. Churchill, London, p. 241.
- Kiang, N. Y. S., Watanabe, T., Thomas, E. C. and Clark, L. F. (1965). Discharge patterns of single fibers in the cat's auditory nerve, *Res. Monogr. nr. 35*, M.I.T. Press, Cambridge, MA.
- Kim, D. O., Chang, S. O. and Sirianni, J. G. (1990). A population study of auditory-nerve fibers in unanesthetized decerebrate cats: Response to pure tones, *J. Acoust. Soc. Am.* **87**: 1648–55.
- Kim, D. O. and Parham, K. (1991). Auditory nerve spatial encoding of high-frequency pure tones: Population response profiles derived from d' measure associated with nearby places along the cochlea, *Hear. Res.* **52**: 167–80.
- Kistler, W. M., Gerstner, W. and van Hemmen, J. L. (1997). Reduction of the Hodgkin–Huxley equations to a single-variable threshold model, *Neural Computation* **9**(5): 1015–1045.

- Kumar, A. R. and Johnson, D. H. (1993). Analyzing and modeling fractal intensity point processes, *J. Acoust. Soc. Am.* **93**: 3365–73.
- Langner, G. (1992). Periodicity coding in the auditory system, *Hear. Res.* **60**: 115–142.
- Lecar, H. and Nossal, R. (1971). Theory of threshold fluctuations in nerves, *Biophysical Journal* **11**: 1048–1067.
- Levitt, H. (1970). Transformed up-down methods in psychoacoustics, *J. Acoust. Soc. Am.* **49**(2): 467–477.
- Liberman, M. C. (1978). Auditory nerve response from cats raised in a low noise chamber, *J. Acoust. Soc. Am.* **63**: 442–455.
- Liberman, M. C. and Oliver, M. E. (1984). Morphometry of intracellularly labeled neurons of the auditory nerve: Correlations with functional properties, *J. Comp. Neuro.* **223**: 163–176.
- Licklider, J. C. (1954). ‘Periodicity’ pitch and ‘Place’ pitch, *J. Acoust. Soc. Am.* **26**: 945.
- Manis, P. B. and Marx, S. O. (1991). Outward currents in isolated ventral cochlear nucleus neurons, *J. Neurosci.* **11**(9): 2865–2880.
- Mark, K. E. and Miller, M. I. (1992). Bayesian model selection and minimum description length estimation of auditory-nerve discharge rates, *J. Acoust. Soc. Am.* **91**(2): 989–1002.
- McKay, C. M. and McDermott, H. J. (1993). Perceptual performance of subjects with cochlear implants using the Spectral Maxima Sound Processor (SMSP) and the Mini Speech Processor (MSP), *Ear & Hearing* **14**(5): 350–367.
- McKay, C. M. and McDermott, H. J. (1997). Loudness perception with pulsatile electrical stimulation: The effect of interpulse intervals, *Submitted to J. Acoust. Soc. Am.*
- McKay, C. M., McDermott, H. J. and Clark, G. M. (1995). Loudness summation for two channels of stimulation in cochlear implants: Effects of spatial and temporal separation, *Ann. Otol. Rhinol. Laryngol.* **104**(Suppl. 166): 230–233.
- McKay, C., McDermott, H., Vandali, A. and Clark, G. (1991). Preliminary results with a six spectral maxima sound processor for the University of Melbourne/Nucleus multiple-electrode cochlear implant, *J. Otolaryng. Soc. Austral.* **6**(5): 354–359.
- Meddis, R., Hewitt, M. J. and Shackleton, T. M. (1990). Implementation details of a computation model of the inner hair-cell/auditory-nerve synapse, *J. Acoust. Soc. Am.* **87**: 1813–6.

- Merzenich, M. M. and White, M. W. (1977). Cochlear implants: The interface problem, in F. T. Hambrecht and J. B. Reswick (eds), *Functional Electrical Stimulation: Applications in Neural Protheses*, Vol. 3, Marcel Dekker, Inc., pp. 321–340.
- Miller, C. A., Abbas, P. J., Rubinstein, J. T. and Matsuoka, A. J. (1997). The neurophysiological effects of simulated auditory prosthesis stimulation, *Third Quarterly Progress Report - NO1-DC-6-2111*, Department of Otolaryngology - Head and Neck Surgery and Department of Speech Pathology and Audiology, University of Iowa, Iowa City, IA 52242, USA.
- Miller, C. A., Woodruff, K. E. and Pfungst, B. E. (1995). Functional responses from guinea pigs with cochlear implants. I. Electrophysiological and psychophysical measures, *Hear. Res.* **92**(1–2): 85–99.
- Miller, M. I. (1985). Algorithms for removing recovery-related distortion from auditory-nerve discharge patterns, *J. Acoust. Soc. Am.* **77**: 1452–64.
- Miller, M. I., Barta, P. E. and Sachs, M. B. (1987). Strategies for the representation of a tone in background noise in the temporal aspects of the discharge patterns of auditory-nerve fibers, *J. Acoust. Soc. Am.* **81**: 665–79.
- Miller, M. I. and Mark, K. E. (1992). A statistical study of cochlear nerve discharge patterns in response to complex speech stimuli, *J. Acoust. Soc. Am.* **92**(1): 202–9.
- Miller, M. I. and Sachs, M. B. (1984). Representation of voice pitch in discharge patterns of auditory-nerve fibers, *Hear. Res.* **14**: 257–279.
- Miller, R. K. (1971). *Nonlinear Volterra Integral Equations*, Mathematics Lecture Note Series, W. A. Benjamin Inc., Menlo Park, California. Library of Congress Catalog Card Number 78-173836.
- Montney, L. M., Morris, D. J., Bledsoe, S. C. and Pfungst, B. E. (1998). The relationship between guinea pig psychophysical and single-unit (ICC) strength-duration functions for pulsatile electrical, *ARO Midwinter Meeting Abstracts*.
- Moore, B. C. J. (1977). *Introduction to the Psychology of Hearing*, University Park Press, Baltimore.
- Moore, B. C. J. and Glasberg, B. R. (1996). A revision of Zwicker's loudness model, *Acustica* **82**(2): 335–345.
- Moss, F. and Wiesenfeld, K. (1995). The benefits of background noise, *Scientific American* **273**(2): 66–69.

- Müller, J. W. (1973). Dead-time problems, *Nuclear Instruments and Methods* **112**: 47–57.
- Neely, S. T. (1981). Fourth-order partition dynamics for a two dimensional model of the cochlea, D.Sc. Dissertation, Sever Institute of Washington University, Saint Louis, Missouri.
- Neely, S. T. and Kim, D. O. (1986). A model for active elements in cochlear biomechanics, *J. Acoust. Soc. Am.* **79**: 1472–80.
- Nelson, D. A., Schmitz, J. L., Donaldson, G. S., Viemeister, N. F. and Javel, E. (1996). Intensity discrimination as a function of stimulus level with electric stimulation, *J. Acoust. Soc. Am.* **100**(4): 2393–2414.
- O’Leary, S. J., Black, R. C. and Clark, G. M. (1985). Current distributions in the cat cochlea, *Hear. Res.* **18**: 273–281.
- O’Leary, S. J., Clark, G. M. and Tong, Y. C. (1995). Model of discharge rate from auditory nerve fibers responding to electrical stimulation of the cochlea: Identification of cues for current and time-interval coding, *Ann. Otol. Rhinol. Laryngol.* **104**(Suppl. 166): 121–123.
- O’Leary, S. J., Irlicht, L. S., Bruce, I. C., White, M. W. and Clark, G. M. (1998). Prediction and effect of variance in neural response to cochlear implant stimulation, *In Preparation* .
- Oxenham, A. J. and Moore, B. C. J. (1994). Modeling the additivity of nonsimultaneous masking, *Hear. Res.* **80**: 105–118.
- Paolini, A. G., Clark, G. M. and Burkitt, A. N. (1997). Intracellular responses of rat cochlear nucleus to sound and its role in temporal coding, *NeuroReport* **8**(15): 3415–3422.
- Papoulis, A. (1991). *Probability, Random Variables, and Stochastic Processes*, third edn, McGraw-Hill International Editions, Singapore.
- Parkins, C. W. (1989). Temporal response patterns of auditory nerve fibers to electrical stimulation in deafened squirrel monkeys, *Hear. Res.* **41**: 137–168.
- Perkel, D., Gerstein, G. and Moore, G. (1967). Neuronal spike trains and stochastic point processes. I. The single spike train, *Biophys. J.* **7**: 391–418.
- Pfingst, B. E. (1988). Comparisons of psychophysical and neurophysiological studies of cochlear implants, *Hear. Res.* **34**: 243–252.
- Pfingst, B. E. (1990). Psychophysical constraints on biophysical/neural models of threshold, *in* J. M. Miller and F. A. Spelman (eds), *Cochlear Implants - Models of the Electrically Stimulated Ear*, Springer-Verlag, New York, pp. 161–185.

- Pfingst, B. E., Haan, D. R. D. and Holloway, L. A. (1991). Stimulus features affecting psychophysical detection thresholds for electrical stimulation of the cochlea. I: Phase duration and stimulus duration, *J. Acoust. Soc. Am.* **90**(4): 1857–1866.
- Pfingst, B. E., Miller, A. L., Morris, D. J., Zwolan, T. A., Spelman, F. A. and Clopton, B. M. (1995). Effects of electrical current configuration on stimulus detection, *Ann. Otol. Rhinol. Laryngol.* **104**(Suppl. 166): 127–131.
- Pfingst, B. E., Morris, D. J. and Miller, A. L. (1995). Effects of electrode configuration on threshold functions for electrical stimulation of the cochlea, *Hear. Res.* **85**: 76–84.
- Pickles, J. O. (1982). *An Introduction to the Physiology of Hearing*, Academic Press Inc. Ltd, London.
- Prucnal, P. R. and Teich, M. C. (1983). Refractory effects in neural counting processes with exponentially decaying rates, *IEEE Trans. Syst. Man Cybern.* **SMC-13**: 676–83.
- Ranck, Jr., J. B. (1975). Which elements are excited in electrical stimulation of mammalian central nervous system: A review, *Brain Research* **98**: 417–440.
- Rhode, W. S. and Smith, P. H. (1986). Encoding timing and intensity in the ventral cochlear nucleus of the cat, *J. Neurophysiol.* **56**: 261–286.
- Rose, J. E., Brugge, J. F., Anderson, D. J. and Hind, J. E. (1967). Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey, *J. Neurophysiol.* **30**: 769–793.
- Rothman, J. S. and Young, E. D. (1996). Enhancement of neural synchronization in computational models of ventral cochlear nucleus bushy cells, *Auditory Neurosci.* **2**: 47–62.
- Rothman, J. S., Young, E. D. and Manis, P. B. (1993). Convergence of auditory nerve fibers onto bushy cells in the ventral cochlear nucleus: Implications of a computational model, *J. Neurophysiol.* **70**: 2562–2583.
- Rubinstein, J. T. (1995). Threshold fluctuations in an N sodium channel model of the node of Ranvier, *Biophysical Journal* **68**: 779–785.
- Rubinstein, J. T., Matsuoka, A. J., Abbas, P. J. and Miller, C. A. (1997). The neurophysiological effects of simulated auditory prosthesis stimulation, *Second Quarterly Progress Report - NO1-DC-6-2111*, Department of Otolaryngology - Head and Neck Surgery and Department of Speech Pathology and Audiology, University of Iowa, Iowa City, IA 52242, USA.

- Rubinstein, J. T., Matsuoka, A. J., Miller, C. A. and Abbas, P. J. (1997). Computational model of the auditory nerve: Interesting aspects of the recovery process, *ARO Midwinter Meeting Abstracts*.
- Sachs, M. B., Winslow, R. L. and Sokolowski, B. H. A. (1989). A computational model for rate-level functions from cat auditory-nerve fibers, *Hear. Res.* **41**: 61–70.
- Saleh, B. E. A. and Teich, M. C. (1982). Multiplied-Poisson noise in pulse, particle, and photon detection, *Proc. IEEE* **70**: 229–45.
- Sewell, W. F. (1996). Neurotransmitters and synaptic transmission, in P. Dallos, A. N. Popper and R. R. Fay (eds), *The Cochlea*, Springer, New York, chapter 9, pp. 503–533.
- Shamma, S. A. (1985). Speech processing in the auditory system II: Lateral inhibition and the central processing of speech evoked activity in the auditory nerve, *J. Acoust. Soc. Am.* **78**(5): 1622–32.
- Shannon, R. V. (1983a). Multichannel electrical stimulation of the auditory nerve in man. I. Basic psychophysics, *Hear. Res.* **11**: 157–189.
- Shannon, R. V. (1983b). Multichannel electrical stimulation of the auditory nerve in man. II. Channel interaction, *Hear. Res.* **12**: 1–16.
- Shannon, R. V. (1989). A model of threshold for pulsatile electrical stimulation of cochlear implants, *Hear. Res.* **40**: 197–204.
- Shannon, R. V. (1993). Psychophysics, in R. S. Tyler (ed.), *Cochlear Implants: Audiological Foundations*, Singular Publishing Group, Inc., San Diego, California, pp. 357–388.
- Shepherd, R. K. and Javel, E. (1997). Electrical stimulation of the auditory nerve: I. Correlation of physiological responses with cochlear status, *Hear. Res.* **108**: 112–144.
- Shi, Y. and Hecox, K. E. (1991). Nonlinear system identification by m-pulse sequences: Application to brainstem auditory evoked responses, *IEEE Trans. Biomed. Eng.* **38**(9): 834–845.
- Siebert, W. M. (1970). Frequency discrimination in the auditory system: Place or periodicity mechanisms?, *Proc. IEEE* **58**: 723–30.
- Sigworth, F. J. (1980). The variance of sodium current fluctuations at the node of Ranvier, *J. Physiol.* **307**: 97–129.
- Smith, D. W. (1996). Determinants of perception with a cochlear implant, Funded Grant Application (Department of Health and Human Services, Public Health Service), Duke University, Durham, NC, USA.

- Smith, D. W. and Finley, C. C. (1997). Effects of electrode configuration on psychophysical strength-duration functions for single biphasic electrical stimuli in cats, *J. Acoust. Soc. Am.* **102**(4): 2228–2237.
- Smith, P. H. and Rhode, W. S. (1987). Characterization of HRP-labeled globular bushy cells in the cat anteroventral cochlear nucleus, *J. Comp. Neurol.* **266**: 360–375.
- Snyder, D. L. and Miller, M. I. (1991). *Random Point Processes in Time and Space*, Springer Texts in Electrical Engineering, Springer-Verlag, New York. ISBN 0-387-97577-2.
- Spoendlin, H. and Schrott, A. (1989). Analysis of the human auditory nerve, *Hear. Res.* **43**: 25–38.
- Srulovicz, P. and Goldstein, J. L. (1983). A central spectrum model: A synthesis of auditory-nerve timing and place cues in monaural communication of frequency spectrum, *J. Acoust. Soc. Am.* **73**(4): 1266–1276.
- Stypulkowski, P. H. and van den Honert, C. (1984). Physiological properties of the electrically stimulated auditory nerve. I. Compound action potential recordings, *Hear. Res.* **14**: 205–223.
- Teich, M. C. and Khanna, S. M. (1985). Pulse-number distribution for the neural spike train in the cat's auditory nerve, *J. Acoust. Soc. Am.* **77**: 1110–28.
- Teich, M. C. and Lachs, G. (1979). A neural-counting model incorporating refractoriness and spread of excitation. I. Application to intensity discrimination, *J. Acoust. Soc. Am.* **66**: 1738–49.
- ten Kate, J. H. and van Bekkum, M. F. (1988). Synchrony-dependent autocorrelation in eighth-nerve-fiber responses to rippled noise, *J. Acoust. Soc. Am.* **84**: 2504–13.
- Tong, Y. C. and Clark, G. M. (1986). Loudness summation, masking, and temporal interaction for sensations produced by electric stimulation of two sites in the human cochlea, *J. Acoust. Soc. Am.* **79**(6): 1958–1966.
- Tykocinski, M., Shepherd, R. K. and Clark, G. M. (1995a). Acute effects of high-rate stimulation on auditory nerve function in guinea pigs, *Ann. Otol. Rhinol. Laryngol.* **104**(Suppl. 166): 71–74.
- Tykocinski, M., Shepherd, R. K. and Clark, G. M. (1995b). Electrophysiological effects following acute intracochlear direct current stimulation of the guinea pig cochlear, *Ann. Otol. Rhinol. Laryngol.* **104**(Suppl. 166): 68–71.
- Tykocinski, M., Shepherd, R. K. and Clark, G. M. (1995c). Reduction in excitability of the auditory nerve following electrical stimulation at high stimulus rates, *Hear. Res.* **88**: 124–142.
- van den Honert, C. and Stypulkowski, P. H. (1984). Physiological properties of the electrically stimulated auditory nerve. II Single fiber recordings, *Hear. Res.* **14**: 225–243.

- van den Honert, C. and Stypulkowski, P. H. (1987a). Single fiber mapping of spatial excitation patterns in the electrically stimulated auditory nerve, *Hear. Res.* **29**: 195–206.
- van den Honert, C. and Stypulkowski, P. H. (1987b). Temporal response patterns of single auditory nerve fibers elicited by periodic electrical stimuli, *Hear. Res.* **29**: 207–222.
- van Stokkum, I. H. M. and Gielen, C. C. A. M. (1989). A model for the peripheral auditory nervous system of the grassfrog, *Hear. Res.* **41**: 71–85.
- Vandali, A. E., Harrison, J. M., Huigen, J. M., Plant, K., and Clark, G. M. (1995). Multichannel cochlear implant speech processing: Further variations of the spectral maxima sound processor strategy, *Ann. Otol. Rhinol. Laryngol.* **104**(Suppl. 166): 378–381.
- Vannucci, G. and Teich, M. C. (1978). Effects of rate variation on the counting statistics of dead-time-modified Poisson processes, *Opt. Commun.* **25**: 267–72.
- Vannucci, G. and Teich, M. C. (1981). Dead-time-modified photocount mean and variance for chaotic radiation, *J. Opt. Soc. Am.* **71**: 164–70.
- Verveen, A. A. (1960). On the fluctuation of threshold of the nerve fibre, in D. B. Tower and J. P. Schadé (eds), *Structure and Function of The Cerebral Cortex*, Elsevier, Amsterdam, The Netherlands, pp. 282–288.
- Verveen, A. A. (1961). *Fluctuation in Excitability*, PhD thesis, University of Amsterdam.
- Verveen, A. A. (1962). Axon diameter and fluctuation in excitability, *Acta morph. neerl.-scand.* **5**: 79–85.
- Verveen, A. A. and Derksen, H. E. (1965). Fluctuations in membrane potential of axons and the problem of coding, *Kybernetik* **2**(4): 152–160.
- Verveen, A. A. and Derksen, H. E. (1968). Fluctuation phenomena in nerve membrane, *Proc. IEEE* **56**(6): 906–916.
- Viemeister, N. F. (1983). Auditory intensity discrimination at high frequencies in the presence of noise, *Science* **221**: 1206–8.
- Viemeister, N. F. (1988). Intensity coding and the dynamic range problem, *Hear. Res.* **34**: 267–74.
- Wever, E. G. (1949). *Theory of Hearing*, John Wiley and Sons, Inc., New York.
- White, M. W. (1978). *Design Considerations of a Prosthesis for the Profoundly Deaf*, PhD thesis, University of California, Berkeley.

- White, M. W. (1984a). The multichannel cochlear prosthesis: Channel interactions, *Proceedings of the Sixth Annual Conference of the IEEE Engineering in Medicine and Biology Society "Frontiers of Engineering and Computing in Health Care"*, Los Angeles, California, pp. 396–400.
- White, M. W. (1984b). Psychophysical and neurophysiological considerations in the design of a cochlear prosthesis, *Audiol. Ital.* **1**(2): 77–117.
- White, M. W. (1987). Neurophysiological and psychophysical considerations in the design of a cochlear prosthesis, *Ann. Otol. Rhinol. Laryngol.* **96**(Suppl. 128): 42.
- White, M. W., Finley, C. C. and Wilson, B. S. (1987). Electrical stimulation model of the auditory nerve: Stochastic response characteristics, *Proceedings of the Ninth Annual Conference of the IEEE Engineering in Medicine and Biology Society*, Boston, Ma., pp. 1906–1907.
- White, M. W., Merzenich, M. M. and Gardi, J. N. (1984). Multichannel cochlear implants: Channel interactions and processor design, *Arch. Otolaryngol.* **110**: 493–501.
- Wilson, B. S., Finley, C. C., Lawson, D. T., Wolford, R. D., Eddington, D. K. and Rabinowitz, W. M. (1991). Better speech recognition with cochlear implants, *Nature* **352**: 236–238.
- Wilson, B. S., Finley, C. C., Lawson, D. T., Wolford, R. D. and Zerbi, M. (1993). Design and evaluation of a Continuous Interleaved Sampling (CIS) processing strategy for multichannel cochlear implants, *J. Rehabil. Res. Dev.* **30**(1): 110–116.
- Winter, I. M. and Palmer, A. R. (1991). Intensity coding in low-frequency auditory-nerve fibers of the guinea pig, *J. Acoust. Soc. Am.* **90**(4): 1958–67.
- Winter, I. M., Robertson, D. and Yates, G. K. (1990). Diversity of characteristic frequency rate-intensity functions in guinea pig auditory nerve fibres, *Hear. Res.* **45**: 191–202.
- Xu, J., Shepherd, R. K., Millard, R. E. and Clark, G. M. (1997). Chronic electrical stimulation of the auditory nerve at high stimulus rates: A physiological and histopathological study, *Hear. Res.* **105**: 1–29.
- Young, E. D. and Sachs, M. B. (1979). Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory nerve fibers, *J. Acoust. Soc. Am.* **66**: 1381–1403.
- Young, E. D., Shofner, W. P., White, J. A., Robert, J. M. and Voigt, H. F. (1988). Response properties of cochlear nucleus neurons in relationship to physiological mechanisms, in S. Hassler (ed.), *Functions of the Auditory System*, John Wiley and Sons, New York, pp. 277–312.

- Zeng, F. and Shannon, R. V. (1992). Loudness balance between electric and acoustic stimulation, *Hear. Res.* **60**: 231–235.
- Zilberman, Y. (1997). Speed and flexibility, in G. M. Clark (ed.), *Cochlear Implants. XVI World Congress of Otorhinolaryngology Head and Neck Surgery*, Monduzzi, Bologna, pp. 81–84.
- Zwicker, E. (1958). Über psychologische und methodische Grundlagen der Lautheit, *Acustica* **8**: 237–258.
- Zwicker, E. and Fastl, H. (1990). *Psychoacoustics - Facts and Models*, Springer-Verlag, Berlin.
- Zwicker, E. and Scharf, B. (1965). A model of loudness summation, *Psychol. Rev.* **72**: 3–26.