Co-ordinating Behaviours in an Insect Biorobot

Mark Payne

Doctor of Philosophy
Institute of Perception, Action and Behaviour
School of Informatics
University of Edinburgh
2009
Abstract

Certain insect behaviours appear to function as reflexes when tested with restricted sensory stimuli. When more complex sensory situations are considered it becomes apparent that these behaviours are not independent; they must interact within the nervous system of the animal or through feedback. This dissertation describes experiments and modelling undertaken to address the question of how optomotor following interacts with phonotaxis in the cricket.

Paths of crickets walking in response to calling song and optical flow stimuli were recorded on an open-loop trackball and in an arena. The results were used to guide the development of a new model of co-ordination between the auditory and visual systems, implemented on a miniature robot.

Five initial hypotheses were investigated, based on a previous robotic modelling study: inhibition of the optomotor response, modulation of the optomotor gain, chaining of two behavioural subsystems, summation at the motor output, and efference copy. Experiments with the trackball allowed the first three possibilities to be rejected, but summation and efference copy were more difficult to distinguish between. The first evidence of a possible efference copy-type mechanism comes from modified trackball experiments where a closed feedback loop was established for the visual system. Modulation of open-loop behaviour was observed that depended on the sign of previously experienced visual feedback, suggesting adaptation of an internal signal. However the same effect could be explained by sustained activity in the visual system. The second piece of evidence supporting the efference copy theory came from the robot model; when the summation mechanism was implemented the optomotor response caused the robot to over-compensate after turning to sound. This behaviour was not observed from crickets in the arena.

Implementation of an efference copy-based mechanism on the robot is accomplished using a recurrent network of spiking neurons (liquid state machine). It is proposed that the circuitry of the mushroom bodies might permit such computation in insects. It is shown that the neural network can learn to predict and cancel out self-generated optical flow signals in the robot during phonotaxis, whilst reproducing the behaviour seen on the trackball under equivalent conditions.
Acknowledgements

Firstly, thanks must go to my supervisors, Barbara Webb and Berthold Hedwig, for their advice, encouragement and patience throughout this project. I am especially grateful to Berthold and his research group for going out of their way to allow me access to their laboratory facilities, and for making me feel welcome on my visits there.

A number of people provided practical help with this project: Leanne Scott carried out a number of trackball experiments on my behalf, allowing me to concentrate on the robot model at a late stage. Technical work on the robot sensors was carried out by Rob Macgregor, the arena was built by Douglas Howie, and work on the trackball and stimuli was carried out by Glen in Cambridge. Their work and ingenuity was indispensable.

My fellow students and researchers at IPAB and in the wider Informatics community have been a great support with their willingness to talk science, programming and nonsense, and have made completing this project seem a little less daunting that it otherwise would have done. Finally I would like to thank Becca, who at the end of all this deserves a holiday.

This project was funded by the EPSRC / MRC Neuroinformatics DTC
Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

(Mark Payne)
# Table of Contents

1 Introduction
   1.1 A Proposal ................................................. 1
   1.2 The Cricket as a Model Insect ............................... 2
   1.3 A Methodology ............................................ 4
   1.4 Research Aims ........................................... 6
   1.5 Thesis Outline ......................................... 8

2 The Components of Insect Behaviour .............................. 11
   2.1 Introduction ............................................. 11
   2.2 Common Principles ....................................... 12
      2.2.1 Matched Filters ..................................... 12
      2.2.2 A Note on Terminology ................................ 13
      2.2.3 Complex Behaviour from a Complex Environment ...... 14
      2.2.4 Feedback .............................................. 15
      2.2.5 Summary ............................................... 16
   2.3 Strategies for Investigating Cricket Behaviour ............. 17
      2.3.1 Field Studies ......................................... 17
      2.3.2 Free-Walking Experiments ............................. 17
      2.3.3 Open-Loop Experiments ............................... 18
      2.3.4 Walking Compensators ................................ 20
   2.4 Cricket Phonotaxis ........................................ 20
      2.4.1 Calling song .......................................... 21
      2.4.2 Female Selectivity for Calling Song Features ........ 22
      2.4.3 Dynamics of Orientation .............................. 27
   2.5 Visual Behaviour in Crickets and Other Insects ............ 30
      2.5.1 Optomotor Following .................................. 30
      2.5.2 Other Motion Responses .............................. 31
2.5.3 Object Responses .............................................. 32
2.6 Auditory & Visual Interaction ................................. 32
  2.6.1 A Motivating Example .................................... 32
  2.6.2 Summation of Reflexes at the Motor Output .......... 33
  2.6.3 Chained Subsystems ..................................... 35
  2.6.4 Inhibition .................................................. 36
  2.6.5 Modulation of the Optomotor Gain .................... 38
  2.6.6 Efference Copy .......................................... 39
  2.6.7 Summary .................................................. 43
2.7 Behavioural Clues to the Mechanism of Interaction .... 44
  2.7.1 Open-Loop Tests [Böhm et al., 1991] ................. 44
  2.7.2 The Effect of the Five Mechanisms in Open Loop Tests [Webb and Reeve, 2003] 46
  2.7.3 Summary .................................................. 49
2.8 Detailed Modelling ............................................ 49
  2.8.1 Phonotaxis ............................................... 50
  2.8.2 Insect Visual Behaviour ................................ 51
  2.8.3 Integrating an Optomotor Response .................... 55
  2.8.4 Improving the Match of the Dynamics ................ 58
  2.8.5 Implementing a True Predictive Mechanism .......... 58
2.9 Summary ........................................................ 58

3 The Physical and Neural Substrates of Cricket Behaviour 61
  3.1 Introduction .................................................. 61
  3.2 The Auditory System ....................................... 62
    3.2.1 Directionality of the Peripheral Auditory System [Michelsen et al., 1994] 62
    3.2.2 Auditory Neurons ...................................... 65
  3.3 The Visual System .......................................... 67
    3.3.1 Structure and Function in the Optic Lobe ............ 68
  3.4 Descending Neurons ........................................ 69
    3.4.1 Separation of Modalities .............................. 70
    3.4.2 Maintenance of the Stimulus Structure ............... 70
    3.4.3 Possible Existence of a Population Code ............ 71
    3.4.4 Gating During Standing ............................... 71
7.8 Behaviour Under Random Motor Disturbances ........................... 207
  7.8.1 Results ................................................................. 208
  7.8.2 Predictive Models & Summation Compared .......................... 210
7.9 Discussion ............................................................................. 211
  7.9.1 Fit of the Robot Model to the Cricket Data ......................... 211
  7.9.2 Performance of Predictive Model ...................................... 214
  7.9.3 Implications for Mechanism of Co-ordination ................. 215

8 Summary and Conclusions ......................................................... 219
  8.1 Behavioural Clues to the Mechanism of Interaction .............. 219
  8.2 Insights From the Robotic Model ....................................... 220
  8.3 Mechanism of Combination — Summary ............................... 221
  8.4 The Liquid State Machine as an Analogue for the Mushroom Bodies .. 222
  8.5 The Liquid State Machine as a Substrate for Sensory Predictions .. 223
  8.6 Future Work ................................................................. 224
    8.6.1 Cricket Behaviour ..................................................... 224
    8.6.2 Separating Exafference from Reafference in Learning in the Robot........ 225
    8.6.3 Neural Modelling of Combination .................................. 226
    8.6.4 Adding Further Modalities ......................................... 226
  8.7 Co-ordination of Behaviour ............................................. 226

A Phonotaxis on the Trackball ..................................................... 229
  A.1 Speakers at 90° ......................................................... 230
  A.2 Speakers at 45° .......................................................... 233
  A.3 Response with cylinder in place ...................................... 233
  A.4 Maintenance of phonotaxis over several minutes ............... 234

B Scototaxis in the Arena ............................................................. 237
  B.1 Scototaxis Without Calling Song .................................... 237
  B.2 Influence of Attractive Targets on Phonotaxis ................. 239

C Close-Up Recordings in the Arena ........................................... 241
  C.1 Data Capture ............................................................ 241
  C.2 Tracking ................................................................. 242
  C.3 Separating Behaviour into Bouts ..................................... 243
# List of Figures

2.1 Generic reflex-like pathway .......................................................... 13
2.2 Generic feedback loop ................................................................. 16
2.3 Open loop trackball and walking compensator compared ............... 19
2.4 Elements of cricket calling song ................................................... 22
2.5 Selectivity for syllable rates around 30 Hz .................................. 24
2.6 Auto-correlation for syllable rate recognition ............................. 26
2.7 Split song ..................................................................................... 29
2.8 Pattern frequency dependency of the fly optomotor response .......... 31
2.9 Summation at the motor output ................................................... 34
2.10 “Chaining” of the phonotaxis and optomotor systems ................. 35
2.11 Complete inhibition of optomotor pathway by phonotaxis pathway .. 36
2.12 Adjustment of optomotor gain by phonotaxis pathway ................. 39
2.13 Efference copy from phonotaxis system to optomotor system ....... 40
2.14 Shift of auditory tuning curves with optomotor stimuli ............... 45
2.15 Simulated open-loop multimodal experiments ............................ 46
2.16 Efference copy from the simulation of [Webb and Reeve, 2003] .... 49
2.17 Elementary Motion Detector ......................................................... 53
3.1 Layout of the cricket ear ................................................................. 63
4.1 Cross-sectional view of trackball .................................................. 85
4.2 Trackball installed in sound-proof box .......................................... 86
4.3 Averaged optomotor step responses ............................................. 90
4.4 Example optomotor step response ................................................ 91
4.5 Step response magnitude over twenty minutes ............................ 92
4.6 Stimulus protocol for characteristic curve experiments ............... 93
4.7 Optomotor characteristic curves from ascending / descending sweeps 94
4.8 Averaged optomotor characteristic curve ................................. 94
4.9 Normalised optomotor characteristic curve ........................................ 95
4.10 Optomotor response to sinusoidal inputs of 1 & 2 Hz ............................. 96
4.11 Effect of nonlinear characteristic on sinusoidal inputs ......................... 98
4.12 Optomotor response amplitude / frequency plot .................................. 98
4.13 Optomotor response phase / frequency plot ......................................... 99
4.14 Time-averaged open-loop responses to multimodal stimuli ................... 101
4.15 Illustration of positive & negative feedback conditions ......................... 103
4.16 Feedback test results over chirp period ............................................. 105
4.17 Effect of filtering on standard and split calling song pattern .................. 109
4.18 Directional efference copy mechanism ................................................ 113

5.1 Arena layout ......................................................................................... 119
5.2 Arena with solid and graded stripes ..................................................... 124
5.3 Blank arena tracks ................................................................................ 125
5.4 Striped arena tracks ............................................................................. 126
5.5 Graded arena tracks ............................................................................. 127
5.6 D values for blank & striped conditions ................................................ 128
5.7 Light & I-R arena tracks ...................................................................... 130
5.8 I-R arena tracks with unusual tracks removed ....................................... 131
5.9 D statistic for light and I-R tracks ........................................................ 131

6.1 Heisenberg’s model of olfactory conditioning ........................................ 142
6.2 Khepera robot and additional sensors .................................................. 146
6.3 Simple pressure difference receiver ....................................................... 147
6.4 Ear circuit diagram ................................................................................ 148
6.5 Ear circuit directionality ....................................................................... 149
6.6 Neural controller overview .................................................................... 152
6.7 Dynamic synapse behaviour — slow depression and fast facilitation .... 158
6.8 Dynamic synapse behaviour — fast depression and slow facilitation ...... 158
6.9 Stages of processing in the robot’s visual system .................................... 160
6.10 Characteristic curve for robot optomotor system .................................. 161
6.11 Neural network for auditory sub-system .............................................. 162
6.12 Auditory filter output ........................................................................... 167
6.13 Optomotor response to sinusoidal inputs ............................................. 168
6.14 Output of EMDs ................................................................................... 169
6.15 Output of optomotor system with low-pass filter .................................. 170
6.16 Robot step response ........................................ 171
7.1 Auditory neuron responses to calling song with SRI 18 ms ........ 178
7.2 Auditory neuron responses to calling song with SRI 42 ms .......... 179
7.3 Auditory neuron responses to calling song with SRI 74 ms .......... 180
7.4 Robot locating calling song, phono subsystem only .................. 182
7.5 Effect of optomotor system on biased trajectories .................... 184
7.6 Output of optomotor system reacting to a rotatory bias ................ 184
7.7 LSM response to discrete auditory pulses ............................ 186
7.8 LSM response to calling song during phonotaxis ..................... 187
7.9 LSM output filters ............................................ 188
7.10 Training and testing runs for off-line model fitting .................. 190
7.11 Prediction of LSM with auditory input alone ......................... 191
7.12 Prediction of LSM with auditory and opto input ...................... 192
7.13 Behaviour of auditory- & opto-based prediction with constant bias .. 194
7.14 Evolution of readout weights using Delta Rule ....................... 196
7.15 Fit of prediction from on-line and off-line methods .................. 197
7.16 Fit of non-directional prediction .................................. 198
7.17 Robot locating calling song, summation at motor output ............. 199
7.18 Opto and phono system outputs with opto gain $6 \times 10^{-4}$ ........ 200
7.19 Opto and phono system outputs with opto gain $3 \times 10^{-4}$ ........ 201
7.20 Robot feedback test with no predictive model ....................... 202
7.21 Robot feedback test with directional predictive model — chirp period average ................................................. 203
7.22 Robot feedback test with directional predictive model — example .. 204
7.23 Robot feedback test with non-directional predictive model — chirp period average ................................................. 205
7.24 Robot feedback test with non-directional predictive model - example. 206
7.25 Sound localisation under random disturbances ....................... 209
7.26 D Statistics for robot under random disturbances ................... 210

A.1 Example responses to standard phonotaxis test ...................... 231
A.2 Measuring phonotactic steering .................................... 232
A.3 Distribution of phonotactic steering strengths ......................... 232
A.4 Steering speeds with sound from 45° and 90° ........................ 233
A.5 Phonotaxis responses over eight minutes ............................ 234
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>B.1</td>
<td>Scototaxis examples</td>
<td>238</td>
</tr>
<tr>
<td>B.2</td>
<td>Phonotaxis with attractive target</td>
<td>240</td>
</tr>
<tr>
<td>C.1</td>
<td>Example tracking frame</td>
<td>242</td>
</tr>
<tr>
<td>C.2</td>
<td>Example close-up data</td>
<td>244</td>
</tr>
<tr>
<td>C.3</td>
<td>Mean speed distribution</td>
<td>245</td>
</tr>
<tr>
<td>C.4</td>
<td>Pause interval distribution</td>
<td>246</td>
</tr>
<tr>
<td>C.5</td>
<td>Mean peak speed distribution</td>
<td>246</td>
</tr>
<tr>
<td>C.6</td>
<td>Rotation magnitude vs. speaker angle</td>
<td>247</td>
</tr>
<tr>
<td>C.7</td>
<td>Rotation speed and other parameters vs. speaker angle</td>
<td>248</td>
</tr>
<tr>
<td>D.1</td>
<td>Diagram for Auditory Board</td>
<td>252</td>
</tr>
</tbody>
</table>
List of Abbreviations

$Q_i$ \quad $i^{th}$ quartile

3-D \quad Three-dimensional

A/D Converter \quad Analogue-to-Digital Converter

AC \quad Alternating Current

ADALINE \quad Adaptive Linear Element

AN1 \quad Ascending Neuron 1 of the cricket auditory system

AN2 \quad Ascending Neuron 2 of the cricket auditory system

ANOVA \quad Analysis of Variance

API \quad Application Programming Interface

BN \quad Brain Neuron

BNC1/2 \quad Cricket auditory brain neuron, Class 1/2

CCD \quad Charge Coupled Device

CR \quad Conditioned Response

CRI \quad Chirp Repetition Interval

CS \quad Conditioned Stimulus

D \quad Directedness Statistic

DBNi5 \quad Cricket Descending Brain Neuron of cluster i5

DC \quad Direct Current
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>DCMD</td>
<td>Descending Contralateral Movement Detector neuron of the locust visual system.</td>
</tr>
<tr>
<td>df</td>
<td>Degrees of freedom</td>
</tr>
<tr>
<td>EMD</td>
<td>Elementary Motion Detector</td>
</tr>
<tr>
<td>EOD</td>
<td>Electric Organ Discharge</td>
</tr>
<tr>
<td>EPSP</td>
<td>Excitatory Post-Synaptic Potential</td>
</tr>
<tr>
<td>F–B</td>
<td>Forward–Backward</td>
</tr>
<tr>
<td>GABA</td>
<td>γ-Aminobutyric Acid</td>
</tr>
<tr>
<td>H-R model</td>
<td>Hassenstein-Reichardt model</td>
</tr>
<tr>
<td>H1 ... H4</td>
<td>Identified horizontal cells of the fly lobula plate.</td>
</tr>
<tr>
<td>HSE</td>
<td>Identified ‘equatorial’ cell of the fly lobula plate HS cells.</td>
</tr>
<tr>
<td>I-R</td>
<td>Infra-red</td>
</tr>
<tr>
<td>IPSP</td>
<td>Inhibitory Post-Synaptic Potential</td>
</tr>
<tr>
<td>KC</td>
<td>Kenyon Cell</td>
</tr>
<tr>
<td>L–R</td>
<td>Left–Right</td>
</tr>
<tr>
<td>LGMD</td>
<td>Lobula Giant Movement Detector neuron of the locust visual system</td>
</tr>
<tr>
<td>LSM</td>
<td>Liquid State Machine</td>
</tr>
<tr>
<td>MB</td>
<td>Mushroom Body/Bodies</td>
</tr>
<tr>
<td>Mics</td>
<td>Microphones</td>
</tr>
<tr>
<td>MS</td>
<td>Mean Square</td>
</tr>
<tr>
<td>nMSE</td>
<td>Normalised Mean Square Error</td>
</tr>
<tr>
<td>NS</td>
<td>Not Significant</td>
</tr>
<tr>
<td>ON1</td>
<td>Omega Neuron 1 of the cricket auditory system.</td>
</tr>
<tr>
<td>ON2</td>
<td>Omega Neuron 2 of the cricket auditory system.</td>
</tr>
</tbody>
</table>

xx
Identified extrinsic neuron contacting the penduculus of the honeybee mushroom body. Thought to be involved in conditioning of proboscis extension.

Projection Neuron

Primary Odour Quality

Pearson’s correlation coefficient

Revolutions per minute

Root Mean Square

(Sample) Standard Deviation

Sound Pressure Level

Syllable Repetition Interval

Sum of Squares (in ANOVA)

Spike-Timing-Dependent Plasticity

Unconditioned Stimulus

Visual Sampling Unit

Ventral Unpaired Median cell of Maxillary Neuromere 1

Extrinsic Neuron
Chapter 1

Introduction

Insects have taught us, and continue to teach us, a great deal about animal behaviour. Observations in the field and the laboratory have revealed numerous ways in which insects navigate, construct living places, find mates and reproduce, obtain food, and evade predation. While these activities are essential to life and might therefore be termed basic, they are far from simple. The fact that insects carry them out using brains containing perhaps as few as one hundred thousand neurons, compared to a human’s one hundred billion, has proven thought provoking for ethologists, neuroscientists and psychologists, and inspiring for engineers and practitioners of artificial intelligence.

Over decades of research certain principles have emerged that go some way towards explaining how insects achieve this feat: reflex-like behaviours that trade off speed against accuracy and adaptability, sensors that work as task-specific “matched filters” [Wehner, 1987], and a close interaction with the physical environment that can lead to emergence of systems and structures more complex than the individual insects themselves [Bonabeau, 1999]. Beyond these powerful abstract concepts, however, it is not easy to make generalisations about insect sensorimotor systems. The ecological niches occupied by the many insect species are so diverse that the different behaviours studied often serve only to illustrate a set of equally diverse mechanisms.

1.1 A Proposal

The first proposition of this thesis is that in order to find further generality in the organisation of insect behaviour it is instructive to look at the interactions between these seemingly separate systems. Conventional neuroethological studies are effective at isolating the stimuli and neural pathways necessary to elicit and govern a particular be-
haviour under controlled conditions. In order to avoid confounding effects the sensory modalities that are not essential are excluded. In an animal's usual habitat, by contrast, sensory input is continuous and multimodal. Different stimuli that produce unambiguous responses when presented in isolation will occur together, activating multiple sensory systems. If these simultaneously active sensory pathways are to guide intelligent action then they need to interact in an appropriate way. By looking at mechanisms of interaction there is a good chance of finding generality and structure, because the same kinds of interaction are likely to occur in many situations.

Possible kinds of interaction include choice behaviour, cue integration and association in memory. Choice behaviour refers to situations where the insect responds to only one of a number of simultaneous stimuli, even though each stimulus would result in some action if presented individually [McFarland and Bösser, 1993]. This type of interaction is advantageous in situations where the individual actions would conflict in some way. Conversely, cue integration is appropriate where multiple modalities provide complementary information about a single environmental feature. Combining multimodal signals by taking into account their individual reliabilities is something that the human brain does in certain circumstances [Knill and Pouget, 2004] and insect brains might perform similar combinations [Cheng et al., 2007]. Associative learning is of value to an animal when one stimulus is a predictor of another. In classical conditioning paradigms an initially neutral stimulus becomes associated with a reward or a punishment, and in operant conditioning paradigms the predictor is the animal's own action. Examples of both types of learning have already been discovered in insects and other invertebrates [Bitterman et al., 1983] [Carew, 2000a] [Wolf et al., 1992] [Brembs, 2003].

1.2 The Cricket as a Model Insect

The work described in this dissertation concentrates on a particular insect: the cricket *Gryllus bimaculatus*. This cricket, together with a few closely related species, has become a so-called “model organism” for the study of phonotaxis, that is the localisation of sound signals by locomotion towards their source [Carew, 2000b]. Adult female *Gryllus* use phonotaxis to find potential mates, who signal their presence by the production of a stereotypic calling song [Bennet-Clark, 1989]. Females can recognise the carrier frequency and pattern of the song, distinguish its direction and run towards it [Popov and Shuvalov, 1977].
In order to successfully study behavioural interaction it is necessary to be able to reliably elicit and observe a pair of behaviours. For this purpose, phonotaxis is very suitable. The calling song is the only stimulus needed to initiate tracking [Regen, 1913], which may be maintained for an hour or more as long as the sound pattern persists [Weber et al., 1981]. Furthermore, once initiated, the steering of the cricket appears reflexive in nature. A pulse of sound of the right frequency played from either side of the body elicits a steering movement to that side [Poulet and Fledwig, 2005]. These characteristics of phonotaxis make it a convenient choice for study in behavioural experiments, and several types of apparatus have already been designed to facilitate its study.

The behaviour selected for investigation in combination with phonotaxis was the optomotor following response. If a cricket or other insect is placed at the centre of a vertically patterned cylinder, which is then rotated, the insect tends to turn in the direction of rotation [Srinivasan et al., 1999b]. This serves to reduce its perceived optical flow, and stabilise it with respect to its surroundings. In an insect’s natural habitat the optical flow might be caused by the wind blowing it around as it flies, for example. The optomotor response would help it to compensate and fly straight.

Optomotor following is of particular interest in the current context because it presents a classic problem of behavioural interaction. Any movement of an insect in the horizontal plane results in horizontal optical flow at the retina, whether the cause of the movement is an external disturbance, a turn in response to some stimulus such as a chirp of cricket calling song, or a mixture of the two. The optomotor response as described would act to oppose optical flow resulting from all of these situations. The supposed utility of the optomotor response for the insect comes from its tendency to counteract external disturbances, but the tendency to oppose intentional movement is disadvantageous. [Holst, 1973] suggests that the solution is for the insect to separate out the external (exafferent) and self-generated (reafferent) components of its optical flow by predicting the latter based on copies of its motor commands. This is not the only solution however. Alternatives are set out and discussed in Chapter 2.

Responding intelligently to exafferent and reafferent input is a problem encountered by all animals, whenever their sensory input can be affected by their own actions. To give a complex example, in a conversation between two people both parties hear each other’s speech as well as their own. In order to have a meaningful conversation the two types of information have to be treated differently.

Whilst optomotor following has not previously been well characterised in crickets,
numerous studies in other insects suggest that the underlying mechanism is likely to be conserved across species and that the behaviour is likely to be as reliable as phonotaxis [Heisenberg and Wolf, 1988]. This, together with the large volume of research findings already published regarding phonotaxis, and existing modelling studies of both behaviours, meant that the research described here could focus on the problem of behavioural co-ordination without requiring a disproportionate amount of time to be spent on understanding the component behaviours individually.

1.3 A Methodology

A combination of behavioural experimentation and biorobotic modelling is used to investigate the interaction between the two selected behaviours. The purpose of the robot is to act as a test platform for different hypotheses about the mechanism of interaction [Webb, 2000]. The behavioural experiments serve to clarify previous experimental results and to collect new data in order to suggest hypotheses and constrain the parameters of the robotic model.

Conventional neuroethology has tended to neglect the potential contribution of model building. Although researchers undoubtedly have hypotheses in mind when they describe the results of an experiment, there is rarely any intention to instantiate them. There are several reasons to be cautious about modelling biological systems. Animals are complex, and to model a behaviour using a set of differential equations or a simple reactive robot clearly involves abstracting away a lot of detail. Some aspects of an animal's physiology might be better understood than others, so the construction of a complete working model would require assumptions to be made about the functioning of the less well-understood parts. Further inaccuracies would be introduced due to the limitations of the media that are used to construct the model. For example we have no material at our disposal that can emulate the dynamic capabilities of muscles.

By testing hypotheses out on the animals themselves the potential problems that might arise from over-abstraction, incorrect assumptions or restrictive media are avoided. However, there are inherent limitations to this approach. Having identified a behaviour of interest, and proposed a mechanism by which it might arise, how is it possible to be certain that that mechanism is present, and truly sufficient to explain the behaviour? Depending on the complexity of the hypothesised mechanism it may be hard to be sure of its sufficiency, especially if it is stated in a qualitative way. This is illustrated
by a recent modelling study which rejected a theory proposed by Götz, that collision avoidance in *Drosophila* could emerge from low-level properties of motion detection [Götz, 1975]. In its original form Götz's proposal sounded plausible, but a quantitative model, matched to the neural properties and flight kinematics of the fly, showed that the hypothesised effect was insufficient to prevent collisions [Lindemann et al., 2008].

For similar reasons it may be difficult to understand how a proposed mechanism would function in new situations that would help to corroborate the hypothesis. As discussed in Chapter 2, insects are often tested in ways that make it uncertain whether sensorimotor systems are functioning in a normal way. By changing the way an insect is tested the operation of individual systems might change, or completely different sensory pathways might take over. For example, Heisenberg and Wolf conclude that *Drosophila* are capable of distinguishing between open- and closed-loop opto-motor stimulation [Heisenberg and Wolf, 1988] (see also Section 2.3). This is where modelling can complement more traditional animal testing. A model permits an experimenter to take exactly those components which are proposed to be necessary to explain a behaviour, and demonstrate whether or not they are sufficient. The experimental conditions can be changed, and predictions generated.

The previously-mentioned issues of abstraction, accuracy, and restrictive media are important, and must be acknowledged when designing biological models [Webb, 2001]. The level of description and amount of abstraction that is appropriate for a model depends on the level of explanation that is sought for a behaviour, and the amount of detailed knowledge that is available. For the insect behaviours under consideration in this project the models are described in terms of networks of spiking neurons, and algorithmic interactions between them. This is because what is ultimately sought is a neural-level explanation for the cricket's behaviour. While the active properties of neurons are clearly the product of lower-level processes there are good reasons for abstracting them away: at the level of the network the interactions between neurons can be described adequately by descriptive rather than mechanistic modelling; introducing the extra detail would make the model harder to understand and more costly to simulate; our understanding of the underlying cellular processes is incomplete, therefore a mechanistic model would actually be less accurate than a descriptive one.

The use of a physical medium is where biorobotics differs from other biological modelling techniques such as cybernetics or theoretical neuroscience. The key advantage is that while cybernetics and theoretical neuroscience generally consider the behaviour of a system with artificial inputs, a biorobot takes its inputs directly from
the world using its sensors. Similarly, the output of a biorobot is movement in the real world, resulting in real feedback. As discussed in the next chapter, these factors are an essential component of biological behaviour, and it is for this reason that the robotic approach is preferred. The kinds of sensors and effectors available for constructing biorobots are certainly not perfect analogues for insect equivalents. However their use may be justified so long as their limitations are acknowledged in the evaluation of the robot.

In the current project behavioural testing is conducted alongside robotic modelling. This is because biorobotics is intended to be an integral part of the scientific process, not a separate activity. There is a danger that the biological relevance of a model will be lost if there is no recourse to biological validation. As will be discussed, the problem addressed in this thesis had already been approached using the biorobotic method. The previous studies had involved modelling only, and the robot had been developed as far as possible without becoming unduly speculative. In order to progress it was necessary to seek clarification from biology before a more informative model could be built.

1.4 Research Aims

A previous biorobotic modelling study [Webb and Harrison, 2000] [Webb and Reeve, 2003] implemented five distinct mechanisms of interaction between phonotaxis and optomotor following. No one of these mechanisms produced a clearly better match to the target behaviour of the cricket walking on an open-loop trackball. The overall aim of the work described in this thesis was to address the limitations of the study in order to:

- Assess the plausibility of the five hypotheses.
- Refine and re-implement the plausible ones in an improved robotic model that could be used to further distinguish between them.

The first limitation of the previous study concerns the accuracy of the dynamics of the robot's optomotor and auditory systems. Inaccuracies arose primarily from a lack of cricket-specific biological data, which resulted in the need to make assumptions in order to build a functioning robotic system.

Optomotor sensing was accomplished using an analogue vision chip with its temporal parameters tuned to the response observed in the fly Drosophila. Although the
1.4. Research Aims

neural mechanisms underlying optomotor following in the cricket are expected to be similar to those in the fly, the different environments and behaviours of these two species might, at least, have been expected to involve some divergence of the temporal tuning of the system. The Drosophila model was used in the absence of any data that would have allowed the authors of [Webb and Reeve, 2003] to check this possibility.

Steering towards sound in the robot of [Webb and Reeve, 2003] used a scheme in which the instantaneous spike rates in a bilateral pair of model motor neurons determined the rate of turning of their respective wheels. Since the study was published, new information about the steering of crickets over time scales of a few milliseconds has shown that this kind of continuous adjustment of turning direction and speed is a feature of the real insect (see Section 2.4.3). When the robot was constructed, however, it was not possible to match the filtering properties of motor output to those of the insect. It was only possible to introduce a reasonable degree of filtering to produce smooth turning of an appropriate speed. Because the robot’s tracks were a product of continuous interaction between the optomotor and phonotaxis systems these inaccuracies had the potential to affect the observed behaviour significantly.

The second factor limiting the conclusions of the previous study was a lack of uncontroversial data with which to compare the robot’s behaviour. Only one published paper [Böhm et al., 1991] describes experiments explicitly conducted to test the interaction between phonotaxis and optomotor following in crickets. Its results are not incompatible with any of the mechanisms of [Webb and Reeve, 2003], and hence cannot be used to rule any of them out. In part this is due to the low numbers of insects that were tested, which means that the key data (graph reproduced in Figure 2.14) is rather noisy. However, even if the data had been of higher quality, the experiment conducted could not have uniquely identified one of the mechanisms under consideration.

In order to overcome these limitations it was clearly necessary to collect new behavioural data from crickets. Some data would be used to distinguish directly between the previously implemented hypotheses. The rest would be used to help build a new robot controller that would be a better fit to the target animal, and hence have more power as a predictive model.

Whilst it was envisaged that the process of investigation and modelling would be guided primarily by behavioural and functional constraints a further aim was to consider how the mechanisms under investigation might be produced by the neural circuitry of the insect brain. A large proportion of the neurons in a cricket’s brain (roughly one third) belong to areas of glomerular neuropil, which do not lie on unimodal sen-
sorimotor pathways, but seem to be areas of convergence for secondary branches from multiple sensory regions [Homberg, 2005]. These areas, namely the mushroom bodies and the central complex are widely conserved across insect species, and ought to be considered in any search for general mechanisms of behavioural co-ordination.

1.5 Thesis Outline

The rest of the dissertation is split into seven main chapters. The first elaborates on the basic principles of behaviour mentioned in this introduction, and discusses the most important results relating to cricket phonotaxis and visual behaviour. The way in which optomotor following and phonotaxis interact through feedback is explained, and five hypotheses about how the cricket might overcome this interaction are introduced: summation, chained sub-systems, inhibition, gain modulation and efference copy. The expected effects of these mechanisms under open-loop conditions are examined with reference to [Webb and Reeve, 2003]. This provides a basis for interpreting the behavioural test results of Chapters 4 and 5.

To understand the design of the robotic model, and some of the reasoning underlying the conclusions of Chapters 4 and 5, requires an appreciation of the known neurophysiology of the cricket. Chapter 3 reviews this material, and also discusses the properties of the mushroom bodies, a structure of the insect brain that may be involved in processing sensory predictions.

Chapter 4 describes how an open-loop trackball system was used to characterise the dynamics of the cricket optomotor response. It then describes the replication of an experiment that looks at time-averaged turning tendency under combined auditory / visual stimulation. A further test is introduced, that examines the crickets' responses under conditions of altered visual feedback, and reveals that some kind of feedback-dependent adaptation is taking place. The validity of the five hypotheses are considered in the light of the results, leading to a rejection of the inhibition, gain modulation and chained subsystems theories.

Chapter 5 covers a second set of experiments in which crickets were recorded on video while performing phonotaxis under free-walking conditions in an arena. These were conducted to test the suggestion made in [Weber et al., 1981] that crickets walk straighter during phonotaxis in the presence of visual structure. Demonstrating a significant change in the crickets' paths due to manipulation of the visual environment would have supported the additive hypothesis, or the presence of an efference copy
mechanism without adaptation. The implications of failing to find any such change are discussed.

Chapter 6 describes the design and implementation of a robotic model for testing the two preferred hypotheses (summation at the motor output and efference copy). The robot’s controller has three sub-systems: auditory, optomotor and predictive. The auditory and optomotor systems are tuned to match the dynamics revealed by the behavioural experiments. The predictive subsystem is based on a recurrent network of spiking neurons (liquid state machine [Maass and Natschläger, 2002]).

Chapter 7 presents the results of testing the robot under conditions equivalent to those experienced by the crickets on the trackball and in the arena. The robot’s performance suggests that summation at the motor output is not the strategy adopted by the cricket. It confirms the assertion, made following the trackball experiments, that a directional efference copy would not explain the feedback-dependent adaptation seen in the cricket. It does not suggest, however, that the non-directional efference copy proposed as an alternative is any better as an explanation.

The final chapter discusses the findings and implications of the work, including possible mechanisms of learning in the mushroom bodies.
Chapter 2

The Components of Insect Behaviour

2.1 Introduction

This dissertation examines a specific insect — the cricket — in order to try and draw conclusions about a general problem — multimodal co-ordination of behaviour. Beginning with the general this chapter first defines some concepts that underlie the current understanding of insect behaviour. These concepts are useful because they can act as a framework for piecing together a model from disparate pieces of biological evidence. It then proceeds to review the existing literature on the specific behaviours of interest: phonotaxis and optomotor following. A treatment of the relevant information on the neural basis of these behaviours is given in Chapter 3.

Section 2.6 illustrates the paradox of optomotor following that was summarised briefly in the introduction, and outlines a number of schemes that have been proposed by previous authors to account for the lack of a detrimental effect on locomotion. The plausibility of the different schemes is supported with reference to examples from cricket research and other insect species, and a number of possible variations on these models are discussed. Five classes of model are highlighted as alternative hypotheses for behavioural investigation:

1. The auditory and optomotor systems sum their effects at the motor output stage.
2. Auditory steering signals are transmitted through the optomotor system.
3. The auditory system inhibits the optomotor system when calling song is present.
4. The gain of the optomotor system is modulated depending on the auditory situation.
5. A copy of the efferent motor command of the auditory system is used to predict the self-generated optomotor input.

The implications of the different models for cricket behaviour on the open-loop trackball are discussed with reference to previous modelling studies. Finally, these modelling studies are critically appraised, leading to the identification of several aspects for improvement that were addressed by the current study.

## 2.2 Common Principles

The previous chapter briefly mentioned a number of ideas which help explain how the small brains of insects can be sufficient for the tasks they perform. These concepts are useful because they can be used in a first attempt to decompose a sensorimotor system into understandable parts.

### 2.2.1 Matched Filters

It has been noted that many insect sensory systems are task-specific in the information they extract. For example the ear of the cricket, discussed in Section 3.2.1, is of a physical size and layout that means it responds with high directionality to sound at the carrier frequency of the species-specific calling song. Furthermore, certain interneurons at the first stage of the auditory pathway are sharply tuned to this frequency. Sensory systems that perform such an early extraction of behaviourally relevant information are termed “matched filters” by Wehner [Wehner, 1987]. A matched filter coupled with a simple reflex-like connection to motor systems can produce a timely response to an environmental stimulus using low numbers of neurons. From an engineer’s perspective such a system may perform sub-optimally; for example it may only respond accurately to a small range of stimuli. From an evolutionary perspective, however, the system only has to be good enough to cope with the range of stimuli actually experienced by the animal.

Sensorimotor pathways that can be divided into a matched filter and a reflex-like connection to motor action are much in evidence across the insects. Figure 2.1 depicts the general form of a such a pathway in a block diagram. A particular sensory modality is detected by the sensor, and the signal is filtered for relevant features. The feature

---

1. The term matched filter is used in a related but more specific sense in signal processing [Turin, 1960]
2.2. Common Principles

Figure 2.1: Stages of a generic reflex-like pathway with a "matched filter" as input. The grey box delineates the animal.

The extraction block is drawn joined to the sensor to indicate that this filtering may be at least partly a property of the physical sensing process. The filtered signal is transmitted to a motor centre, which transforms it into the required form to activate the relevant muscles to produce a response.

2.2.2 A Note on Terminology

Afference & Efference

The words afferent and efferent, as indicated in Figure 2.1, are used in this work to refer to incoming sensory signals and outgoing motor commands respectively. They are sometimes used more generally, to refer to any part of a pathway that conducts signals towards or away from higher brain centres. In the insect brain networks under consideration this distinction can be a hard one to make, so the more restricted sense of the terms is preferred.

Reflex or Reflex-Like?

Describing a particular stimulus evoked action as a reflex may imply that the response is innate not learnt, or that the response is not modifiable (for example by conscious decision). While the former appears to be true of phonotaxis and optomotor following there is at least some evidence to suggest that both behaviours are affected by what might be termed the “motivation” of the animal. For this reason the non-committal term “reflex-like” is used.
2.2.3 Complex Behaviour from a Complex Environment

Given a mechanism of the type shown in Figure 2.1, what kinds of things can an insect do? The answer depends on the environment. Simple reflex-like nervous pathways can produce complex, continuous behaviour if the relevant environmental signals are also complex and continuous.

Taxes

Phonotaxis is one of a category of behaviours termed taxes that involve orientation and motion either towards or away from a stimulus source (positive or negative taxis respectively) [Fraenkel and Gunn, 1961]. In a bilaterally symmetric organism with spatially separated sensors a reflex which causes turning towards the more strongly stimulated sensor would produce positive taxis, whereas turning away from the more strongly stimulated side would produce negative taxis. In simpler animals with only a single sensor orientation may still be achieved by making alternate sampling movements to the left and right (klinotaxis). Provided with the right sort of matched filter these kinds of behaviour could mediate many of the navigational tasks that an insect has to achieve. The mechanism is simplistic, but it has the advantageous property of being robust against changes in the position of the signal source.

Emergent Behaviour

Only the very simplest organisms have a single stereotyped response of the type described above. Typically an insect will respond to many classes of environmental stimuli. If the environment changes through time then successions of different sensory pathways will become activated, and the behaviour of the insect will change accordingly. If the environment is being actively modified then this can lead to interesting and useful outcomes. Stigmergy is the name given to the process whereby groups of individuals become co-ordinated in their behaviour due to environmental traces such as pheromones or physical structures [Bonabeau, 1999]. This type of interaction was originally recognised and studied in social insects. The best known examples involve nest building, where partially completed nest structures act as triggering stimuli for members of the colony to deposit material in a configuration that continues the construction process, e.g. [Karsai, 1999]. The physical environment effectively acts as a communication channel between the insects, although the actions they perform do not superficially appear to be concerned with communication.
Stigmergy is one type of emergent behaviour, where a collection of relatively simple sub-systems (the individual insects) interact in a way that produces a more complex super-system (the colony and its nest). While the definition of what classes as emergent is a matter for philosophical debate, the term has also been used in mobile robotics to refer to the observed behaviour of a single robot equipped with multiple reflexes. For example, a reflex that causes turning away from obstacles coupled with another that causes movement towards walls can result in wall-following behaviour. A human observer might perceive the wall-following as a single, coherent, goal-directed behaviour, but internally there is no "wall following system". The observed behaviour is deemed to have emerged from the interaction of the two simpler sub-systems.

Summary

These examples emphasise the fact that the behaviour of an insect is just one observable part of a dynamical system comprising the brain, body and environment. If the full complexity of the environment is taken into account then quite simple static reflexes in the brain may suffice to explain its behaviour.

2.2.4 Feedback

As an animal moves through its environment its actions continuously influence its future sensory inputs. Depending on the way these inputs are filtered the behavioural state of the animal may change only sporadically. A feedback loop exists, but it is not necessarily one that is amenable to the type of mathematical analysis traditionally applied to feedback systems. A few behaviours do lend themselves to this type of analysis, and the optomotor response considered throughout the rest of this thesis is a classic example. The environmental quantity that determines the response is pattern frequency across the retina, and the resulting turning response directly changes this quantity after a lag of a few tens of milliseconds. The optomotor response acts like a negative feedback controller with a fixed point of zero pattern frequency.

Figure 2.2 shows a generic feedback loop in block diagram form. A reflex pathway exists as in Figure 2.1, the only difference being that actions of the effector have a causal influence on the signal received by the sensor. Such a self-generated component of sensory input is referred to as reafferenter (after [Holst, 1973]) whereas the independent component is referred to as exafferent. The wind blowing an insect around as it flies, as mentioned in the introduction, is an example of something causing exafferent
input.

The box labelled “Environmental Filtering” emphasises that the feedback signal is likely to be modified in some way during physical transmission. In the case of the optomotor response the essential quantity extracted by the visual system is pattern frequency, which depends on the structure of the visual scene [Srinivasan et al., 1999b]. This effect would be summarised by a multiplicative constant incorporated in the “Environmental Filtering” stage relating the angular velocity of the cricket to pattern frequency at the eyes. This constant would change according to the insect’s current surroundings. The environmental effect could be more complex for other modalities.

2.2.5 Summary

The preceding sections have illustrated a simple reflex pathway structure that can be used as an initial model of the sensorimotor processing underlying a chosen behaviour. There is no straightforward way of moving from such an abstract model to one that could be implemented on a robot. As the examples of stigmergy illustrated, observed behaviour may be a product of a collection of parallel pathways, the number and type of which is unknown. In some systems the details of the feature extraction, motor centre and environmental filtering stages might be best explained as complex block diagrams in their own right. For reflex-like behaviours such as phonotaxis and optomotor following, however, it was not expected that a fundamental change from the feed-forward structure shown would be required.
2.3 Strategies for Investigating Cricket Behaviour

Ethologists investigating cricket phonotaxis and other behaviours have used a range of experimental techniques and apparatuses. These differ in the extent to which the sensory input of the insect can be controlled. Whilst being able to accurately determine the sensory input is desirable, increased control over that input is achieved by restraining the insect in some way. This has the effect of altering proprioceptive feedback or even breaking some feedback loops altogether, potentially leading to artefactual behaviour which is not representative of the animal’s normal state. When attempting to understand the insect as a dynamical system it is important to appreciate what technique was used to collect the experimental data, since data from different set-ups are not always directly comparable.

2.3.1 Field Studies

Field studies offer the most natural conditions for the study of behaviour and the most challenges for the production of well controlled experiments. Physical practicalities can make obtaining detailed and accurate data difficult. Sensory inputs vary much more widely than in laboratory experiments, and behavioural measurements or scores tend to vary accordingly. However, as highlighted in the previous section, the environment is an essential component of natural behaviour. Since neuroethology and biorobotics generally consider behaviours that were originally observed in the field it would be ideal if theories developed about the underlying mechanisms could ultimately be validated back in the wild. A recent study has done this for a stochastic model of cricket orientation [Mhatre and Balakrishnan, 2007], and the robotic experiments of [Reeve and Webb, 2003] (discussed in Section 2.8.1) have been translated to an outdoor setting [Reeve et al., 2005]. Other researchers have taken neurophysiology equipment outside to investigate the performance of cricket auditory neurons in response to calling song with realistic noise and attenuation [Römer, 1993].

2.3.2 Free-Walking Experiments

Beyond field studies, the first step in investigating phonotaxis is to bring crickets into a laboratory and test them in an artificial arena setting so certain parameters of the environment (for example light level, temperature and auditory noise) can be controlled. However other factors (for example the intensity and relative position of a
sound source) vary as the cricket moves around. This makes it difficult to determine exactly what sensory input the cricket is experiencing at a particular instant.

Arena experiments have proved most suited to choice tests. For example, preference for two calling songs of different patterns can be assessed by setting up two speakers and waiting for the cricket to walk to one or the other [Pollack and Hoy, 1979]. Y-mazes have sometimes been used to constrain the task further [Popov and Shuvalov, 1977]. Assessing detailed behaviour is more difficult. With some work it is possible to extract information about the speed and orientation of a freely walking cricket from video recordings [Mhatre and Balakrishnan, 2007]. However in arenas of a reasonable size the cricket will only walk for a few tens of seconds before reaching the edge, so it is time consuming to gather a great volume of data in this way.

2.3.3 Open-Loop Experiments

A common approach to understanding a feedback system in biology is to open the feedback loop, by stopping the motor output from influencing the sensors [Cruse, 2006]. The aim of such experiment is to determine the input–output characteristics of the forward pathway only in a feedback system of the form shown in Figure 2.2. If this characterisation can be achieved with accuracy then a robot equipped with a model of the pathway should be able to reproduce the behaviour seen in closed-loop conditions in the arena. In studies of cricket phonotaxis the feedback loop is typically opened by tethering the cricket, for example by fixing it to a pin or rod with sticky wax (for example [Hedwig and Poulet, 2004]). It is then possible to play sound or other stimuli to the cricket without its leg movements affecting the position of the stimulus source relative to the body. A cricket fixed in this way is referred to as being under open-loop conditions.

To find the output of the forward pathway it is necessary to find an appropriate way to measure the leg movements. The approach taken by a number of experimenters (and this work, see Chapter 4) is to use a small, light trackball that floats on a stream of air (Figure 2.3, left). The mechanical feedback received by a cricket in this situation is not exactly the same as it would receive in an arena, but the mass and upward force of the ball can be matched to try and approximate the forces that a cricket would experience from its own mass and inertia [Weber et al., 1981].

2 Although leg movements do continuously alter the positions of the ears.
The movements of the ball can be measured by an optical device, for example the chip from a computer mouse. A computer mouse chip measures movement along two orthogonal directions. If placed underneath the sphere as shown in the diagram, with one axis aligned with the cricket’s body axis, then the second component will measure left–right movements of the ball. This gives an indication of the cricket’s tendency to turn. This is not the same as rotation; if the cricket were to rotate the ball precisely about its “North–South” axis then the chip would not register any movement. If, however, a rotation is accompanied by some forward translation then the steering direction and magnitude will be registered. Some set-ups have used a forward–backward sensor underneath the ball, and another mounted horizontally on the “equator”. This picks up rotation, but would be unresponsive to precise side-stepping.

Open loop experiments are, at present, the only type from which it has been possible to gather neural data in synchrony with behavioural recordings [Schildberger and Hömer, 1988]. With the body and head of the cricket fixed, glass electrodes can be inserted into the brain. These allow neurons with stimulus- or motor-correlated activity to be identified, as well as a degree of intervention by hyperpolarising or stimulating individual neurons.

![Figure 2.3: The two most common set-ups for recording cricket behavioural data. Left: Open-loop trackball. The cricket is fixed in place above a small air-suspended sphere. Rotations of the sphere are detected using an optical sensor. Right: Walking compensator. The cricket is unrestrained. The camera-to-motor feedback loop adjusts the sphere so the cricket remains centred on top. The motor outputs can also be recorded in order to reconstruct the insect's walking path.](image-url)
2.3.4 Walking Compensators

Walking compensators (Figure 2.3, right) initially appear similar to open-loop trackballs but do not produce comparable data. A large globe (diameter 50 cm in [Weber et al., 1981]) is mechanically suspended in a bearing and the cricket stands on top. A video tracking device monitors its position. When the cricket walks the movement is registered by the tracker and used to control two motors which rotate the sphere and keep the cricket in place. The cricket’s angular heading is not controlled for; it can turn to face in any direction. There is thus a closed feedback loop, but the feedback is not the same as the cricket would experience in the arena since the translational component of its motion has negligible effect. For phonotaxis experiments a degree of extra control over the stimulus situation is achieved, since the sound source is placed at a desired distance from the globe, and the volume at the cricket’s position then does not alter.

The chief advantage of a walking compensator over testing in an arena is an improvement in the possible duration of an experiment. Additionally, tracking is taken care of automatically if the signals fed to the compensatory motors are recorded. These can be used to deduce the cricket’s heading direction over each sampling interval and re-construct a two dimensional path\(^3\). The data retrieved using this method is somewhat limited since rotations are not picked up unless accompanied by translation. Also, the temporal resolution is not high. For example the sampling interval used in [Schmitz et al., 1982] was 500 ms. Although there is no reason why a faster data acquisition device could not be used it would not retrieve much more detail because the motor feedback to the sphere is damped. This is done so that when the cricket is running there is minimal acceleration of the substrate, and therefore minimal influence on the cricket’s proprioception.

2.4 Cricket Phonotaxis

Producing a biologically relevant model of the interaction between phonotaxis and optomotor following requires a careful consideration of the dynamics of the two behaviours. Section 2.2 outlined the expected features of a sensorimotor pathway mediating a reactive behaviour such as phonotaxis, emphasising the likelihood of a close

\(^3\)The 2-D paths produced from walking compensator recordings are not directly comparable to paths walked in an arena, since the stimulus does not have the same spatial variation. The analogous situation would be an arena where the speaker floats along at a fixed distance from the cricket.
link between motor action and sensing, and a task-specific extraction of relevant environmental features.

If the generic reflex diagram of Figure 2.1 is taken as the first stage in producing a model of the response to calling song then the next step is to ask the questions:

- What features of the sound signal are extracted?
- How is the resulting signal transformed into motion?

These questions are quite suited to behavioural investigation, so the literature addressing them is quite extensive, albeit inconclusive and sometimes contradictory. The key results on each point are reviewed below, beginning with a characterisation of the calling song that drives the response.

### 2.4.1 Calling song

Amongst true crickets it is only the males that sing, by opening and closing their forewings. Three types of song can be distinguished, associated with different behavioural states: calling song, aggressive song and courtship song. The song of principal interest to this study is the calling song which is used by a male to broadcast its presence and attract females. In *Gryllus bimaculatus* it has the form sketched in the top part of Figure 2.4. It can be divided up into short pulses of sound, around 20 ms long, termed syllables. The dominant tone within each syllable has a frequency between 4.5 and 5 kHz. “Chirps” of three to five syllables are produced with a repetition interval of around 500 ms, although this interval varies considerably between animals (SD ≈ 100 ms according to [Doherty, 1985]). Within the chirps the pulses are repeated with an interval of around 45 ms (SD ≈ 5 ms), giving a duty cycle of between 40 and 50% (see lower part of Figure 2.4). Both the chirp repetition interval (CRI) and the syllable repetition interval (SRI) are affected by temperature [Doherty, 1985]. The above values are for 20 and 21°C respectively. Doherty measured CRI values ranging from 340 ms at 30°C to 800 ms between 10 and 12°C, and SRI values from 40 ms at 30°C to 60 ms at 15°C. The syllable length and carrier frequency were unaffected.

Other species of cricket have calling songs that differ by varying degrees from that of *Gryllus bimaculatus*. The ones most commonly used in phonotaxis experiments have been *Gryllus campestris*, *Acheta domesticus* and some *Teleogryllus* species. *G. campestris* calling song is very similar. It has a carrier frequency in the same
range as *G. bimaculatus*, typically four syllables per chirp, and a within-chirp syllable frequency commonly stated as 30 Hz (equivalently, SRI = 33 ms) although this is again temperature dependent. The first syllable of a chirp is usually quieter than the rest [Thorson et al., 1982]. In *A. domesticus* the carrier frequency of the calling song is perhaps a little higher (5 kHz [Nelson and Nolen, 1997]) and chirps tend to contain only two or three syllables with SRI from 50–60 ms [Ewing and Hoyle, 1965]. In *Teleogryllus* the song structure is more complex, and consists of repeated chirps interspersed with trills (e.g. *T. oceanicus* [Bailey and Thomson, 1977]).

### 2.4.2 Female Selectivity for Calling Song Features

In many parts of the world crickets live in habitats that contain multiple cricket species and other chirping insects. Females are able to locate conspecifics based on the unique characteristics of their particular calling song [Popov and Shuvalov, 1977]. Knowing what features of the calling song are important for recognition can be used to guide the search for underlying mechanisms. A common approach has been to characterise the calling song of a particular species in terms of the parameters shown in Figure 2.4, then create sets of synthesised songs in which a chosen parameter is varied\(^4\). The synthetic songs are then played back to crickets which are assessed for orientation in an arena,

\(^4\)In general more than one parameter must be altered at once because, for example, increasing the SRI without increasing the syllable length reduces the duty cycle.
2.4. Cricket Phonotaxis

Response to Pure Tones

\textit{G. bimaculatus} females tested in a Y-maze could be induced to respond by very rough sound models, even pure tones of 5 kHz [Popov and Shuvalov, 1977]. It is not completely clear whether the crickets ran towards the pure tone or simply came out of their shelter to explore the maze; no evidence of tracking of pure tones has been shown in the closely-related \textit{G. campestris} [Weber et al., 1981] [Wendler, 1990]. Most investigators agree that at least some part of the song structure is required for tracking to take place in these two species.

Chirp Length

[Popov and Shuvalov, 1977] found that \textit{G. bimaculatus} displayed no strong selectivity for the number of syllables in a chirp if that number was between three and nine, and also reported that they would continue tracking if a model calling song was switched to one with even longer chirps or a continuous trill. In \textit{G. campestris}, by contrast, they noted a strong selectivity for the number of syllables in a chirp, with three or five being much less attractive than four. This contradicts the findings of [Weber et al., 1981], where it was found that \textit{G. campestris} preferred chirps of eight syllables over chirps with four syllables, and [Thorson et al., 1982], where \textit{G. campestris} were seen to track continuous trills, even without any prior experience of the correct calling song. These results led Thorson and colleagues to put forward a hypothesis that a 30 Hz modulation of the 5 kHz carrier frequency is the necessary and sufficient quantity for eliciting recognition and tracking in \textit{G. campestris}. Based on this hypothesis they predicted and confirmed that the \textit{G. campestris} females they were testing would perform phonotaxis to both the aggressive song of their own species and the very similar calling song of \textit{G. bimaculatus}. This disagreement with Popov and colleagues’ results may reflect the difference in behavioural paradigm, or even differences between geographically separated populations of crickets.

Syllable Rate Within Chirps

Crickets preferentially perform phonotaxis to songs matching the species-specific syllable rate. This has been shown most thoroughly and conclusively by [Thorson et al., 1982] for \textit{G. campestris} and [Doherty, 1985] for \textit{G. bimaculatus}. Thor-
son's results (reproduced in Figure 2.5) show a band-pass response to syllable rate which saturates at 100% around the 30 Hz syllable frequency hypothesised to be essential. Doherty's plot for *G. bimaculatus* is very similar.

![Figure 2.5: Selectivity for SRIs from 15 to 100 ms at 21–22°C. Different symbols denote different crickets. Reproduced with kind permission from Springer Science+Business Media: Journal of Comparative Physiology A, Auditory Behaviour of the Cricket, Vol. 146(3), 1982, Page 368, John Thorson, Theo Weber and Franz Huber, Figure 5.](image)

**Duty Cycle**

[Thorson et al., 1982] adjusted the duty cycle of the standard calling song pattern up to nearly 90% and down to 0.007% whilst still retaining tracking. Reducing the gap between syllables to less than 5 ms, however, sharply reduced phonotaxis.

**Chirp Rate and Trade-Off Phenomena**

Several works have stated that the chirp rate has an influence on phonotaxis. [Popov and Shuvalov, 1977] described chirp repetition interval as a motivational parameter; the CRI did not need to fall within tight boundaries to initiate phonotaxis, but a shorter CRI (within the biologically observed range) was found to be more attractive. [Stout et al., 1983], meanwhile, found that female *Acheta domesticus* approached a synthetic calling song most accurately when the CRI matched the modal value for.
natural male songs. Similarly, [Doherty, 1985] showed that the range of SRIs that were tracked by *G. bimaculatus* was broader when the CRI was at the lower end of the biological range.

More generally [Stout et al., 1983] and [Doherty, 1985] both conclude that the strength of phonotaxis is affected by CRI, syllable duration, and chirp length, and that the overall performance of the cricket is determined by a combination of these factors. Doherty points out that these “trade-off” phenomena may only appear when certain parameters of the song take marginally attractive values. In that case tests such as that described above, for the effect of changing the duty cycle, may fail to reveal an effect if the other parameters of the song are kept near the optimum.

[Wendler, 1990] produced simple sound patterns by sinusoidally modulating the envelope of a 5 kHz tone with frequencies from 0.5 to 100 Hz. These were presented to *G. bimaculatus* on a walking compensator and produced tracking in response to 3 Hz modulations as well as the expected 30 Hz. This disproves the necessity of the 30 Hz syllable structure for eliciting tracking, and suggests that the cricket explicitly filters for the chirp structure.

**Implications for Mechanism of Recognition**

The search for mechanisms underlying pattern recognition in phonotaxis has focused on the uncontroversial data regarding the dependence of tracking on syllable repetition interval. Several mechanisms for recognising a particular SRI have been put forward: auto-correlation, paired high- and low-pass filters, and under-damped neural oscillations [Bush and Schul, 2006]. For *Gryllus* the paired high- and low-pass filter hypothesis fits the behavioural data best, because it predicts a monotonic decrease in response for SRIs lower or higher than the species-specific value. The other two mechanisms predict secondary regions of SRI preference.

The auto-correlation principle [Weber and Thorson, 1989], illustrated in Figure 2.6, supposes that the incoming sound pattern is transmitted through two parallel pathways, one of which contains a pure time delay equalling the desired SRI. The pathways re-converge on some mechanism which calculates the correlation between the direct and the delayed signals. As the figure shows, this would produce a response to the desired SRI, but also to half the desired SRI (as well as one third, and so on). This possibility was considered by [Thorson et al., 1982], and they state that if any secondary region of selectivity was present then their method was not capable of detecting it.

Detection of a secondary region of SRI preference should have been possible had it
Figure 2.6: Auto-correlation for syllable rate recognition. Top — When the SRI of the incoming sound matches the delay the direct and delayed pulses match precisely. Middle — Doubling the incoming SRI results in no coincidences. Bottom — Integer divisions of the SRI result in more coincidences per unit time than the desired SRI.
existed, since a similar secondary preference was shown for the bush cricket *Tettigonia cantans* using a walking compensator [Bush and Schul, 2006]. Bush and Schul show that this insect is selective for its species specific SRI as well as SRIs with *twice* that value. This is indicative of a mechanism involving underdamped oscillation of a neural membrane potential, a possibility which they explore in several further experiments. So far there is no evidence of such a phenomenon in *Gryllus* (B. Hedwig, personal communication).

### 2.4.3 Dynamics of Orientation

The dynamics of walking during phonotaxis determine the input that will be received by the visual system of the cricket. Therefore it is important to consider what information is available to constrain the model motor outputs of the robot.

**Timing of Stops**

During phonotaxis crickets do not run continuously but make repeated pauses. There is currently no information which can allow the timing of pauses and bouts in *Gryllus bimaculatus* to be predicted. Walking bouts in *Scapsipedus marginatus* are correlated with the chirps of the calling song [Murphey and Zaretsky, 1972], but this is likely to be peculiar to this species, where chirps occur at a rate of less than one per second. It does not appear to be the case in *Gryllus*; an average of 25 stops per minute were reported for *G. campestris*, whether or not a recorded calling song was present [Schmitz et al., 1982]. In the same study it was tested whether stops were more likely to occur when the angular deviation of the speaker from the cricket’s midline was large. No such correlation was found.

**The Role of Stops**

Based on video analysis of arena experiments some researchers have proposed that turning is carried out immediately following a pause, and that the rest of a walking bout is not affected by the sound direction. This conclusion was reached both for the sparsely chirping *Scapsipedus marginatus* [Murphey and Zaretsky, 1972] and the continuously trilling *Teleogryllus oceanicus* [Bailey and Thomson, 1977]. Conversely, Mhatre and Balakrishnan observe that for *Plebeioeryllus guttiventris*: “During pauses, females usually stood still” and “The change in heading angle occurred gradually during walking bouts, forming curved paths” [Mhatre and Balakrishnan, 2007].
A "spiky" velocity trace, that is one with frequent stops, was associated with better tracking in *Gryllus campestris* [Weber et al., 1981]. To follow up this observation, Weber and colleagues used the output of their walking compensator to gate a speaker so that crickets heard calling song either only during pauses or only during walking. Both types of gating degraded the tracking performance of crickets, but when calling song was only present during walking the performance was particularly bad, suggesting that sound during stops was the least dispensable component. However, *G. campestris* have been shown to respond to a switch in speaker angle without a pause in walking [Wendler et al., 1980], so sound heard during walking clearly has a role in the normal behaviour of this species.

**Scaling of Turns**

Crickets performing phonotaxis on a walking compensator in the absence of visual structure are observed to oscillate around the speaker direction by as much as ±60° [Weber et al., 1981] at a frequency of 0.2 to 0.4 Hz [Schmitz et al., 1982]. This type of movement could provide a significant input to the optomotor response once visual structure is introduced. How might these oscillations arise?

Earlier, arena based, studies concentrated on the analysis of turning at the level of walking bouts. The orientation of the cricket with respect to the speaker was measured at each pause, and the angular size of the turn during the next bout was recorded [Murphey and Zaretsky, 1972] [Bailey and Thomson, 1977]. Both of the cited studies considered whether the size of turn angles was correlated with the speaker angle during the pause. Murphey and Zaretsky claimed to see a bimodal distribution of turn sizes; turns when the speaker was towards the cricket's rear (between 120° and 180° from the midline) were significantly larger (54±22°) than those elsewhere (33±22°). A similar assertion was made for *Teleogryllus oceanicus* by Oldfield [Oldfield, 1980]. Bailey and Thompson, meanwhile, claimed that turns were scaled within the frontal field (angular deviation less than 30°) for *T. oceanicus*, but of a fixed size (30–40°) elsewhere. If turns of a fixed size were a general feature of all crickets then this might cause them to overshoot the direction of the speaker when correcting for small angular deviations.

Investigations of *Gryllus* on a walking compensator and in open-loop conditions have produced some complementary data. On the former apparatus it is possible to draw a characteristic curve for phonotaxis by calculating the relative angle of the speaker at each (fixed) sampling interval and the angular velocity of the cricket at the next. In *Gryllus campestris* this produced a sinusoidal characteristic curve
[Schmitz et al., 1982], with an average turning velocity of zero in response to sound directly in front or behind the animal, and maximal turning for a speaker position of 90° either side. On an open loop trackball [Böhm et al., 1991] or paired treadwheel [Stabel et al., 1989] the characteristic curve is obtained by moving the speaker in steps around the insect. Both the cited studies confirmed the sinusoidal characteristic curve, suggesting that *Gryllus* do scale their turning.

All the points on the sinusoidal characteristic curves were, however, likely to be averages of multiple turns. This means that they are not necessarily inconsistent with the fixed-size turn idea. The early arena studies indicated that when stopping and turning crickets made a turn to the correct side in around 80% of cases overall. It would be possible for a sinusoid to emerge from the average of many turns if the speaker position influenced the probability of turning to the correct side, with the probabilities being equal at the front and rear, and biased further to the correct side for more lateral positions.

**Rapid Steering**

Examination of the steering of *Gryllus bimaculatus* at high temporal resolution using an open-loop trackball system has revealed that crickets exposed to a “split” song (Figure 2.7) will steer to individual syllables coming from alternate directions [Hedwig and Poulet, 2004]. This suggests that they could employ continuous correction during walking, an idea seemingly at odds with the concept of fixed-size turns seen in other species in the arena. It was also observed that the turns to these pulses occurred with a short latency (55–60 ms) incompatible with the idea that the syllable pattern of a chirp is evaluated as a pre-requisite to turning.

![Figure 2.7](image_url)

Figure 2.7: Illustration of split song principle. Two traces show amplitude modulation of the carrier frequency for left and right channels. The syllable rate on each channel is set so the combined pattern has the SRI of the natural calling song.
Further tests showed that turns to individual pulses were gradually increased in magnitude during the first few chirps of a correct song, after which the crickets would also respond to ordinarily unattractive chirps and pulses inserted into the chirp sequence [Poulet and Hedwig, 2005]. This led Poulet and Hedwig to hypothesise that a parallel pattern recognition process might act to modulate a low-level reflexive steering process.

2.5 Visual Behaviour in Crickets and Other Insects

2.5.1 Optomotor Following

As outlined in the previous chapter, optomotor following is the tendency of an insect or other animal to respond to rotation of its visual field by turning its eyes, head, or body in the direction of the rotation. Under conditions of free movement this turning would reduce the speed of optical flow at the retina and stabilise the eyes with respect to the surroundings. In the cricket, since the eyes are fixed in the head capsule, it is the head or the whole body that has to be rotated to achieve this. The same type of response appears to be common to all insects, although only a very little information has been published about the behaviour in crickets specifically [Böhm et al., 1991]. The one experiment in that study is enough to confirm the existence of optomotor following in the cricket, but it does not provide any quantitative information about the dynamics of the behaviour.

In contrast, study of the response in flies has been extensive [Egelhaaf, 2005]. Experiments in which vertically striped cylinders were rotated at different speeds around tethered flies (i.e. open-loop conditions) have revealed that the motor output is dependent on the temporal frequency of the stripes passing the eye. This is in turn dependent on the spatial frequencies present in the retinal image as well as the speed of rotation [Srinivasan et al., 1999b]. An early model of motion detection [Hassenstein and Reichardt, 1956] based on correlation of the intensity signal at neighbouring points on the retina has proven to be highly predictive of the behaviour of flies in response to these sorts of stimuli (see also Section 2.8.2). It produces a directionally sensitive response with the correct log-Gaussian-shaped characteristic curve as sketched in Figure 2.8.

As pattern frequency increases from zero up to a certain level the response of a fly will also increase. For higher pattern frequencies the fly’s response levels out and starts
2.5. Visual Behaviour in Crickets and Other Insects

![Graph showing pattern frequency dependence of the fly optomotor response](image)

Figure 2.8: Sketch of the pattern frequency dependence of the fly optomotor response to decline again. This reduction in response magnitude at high pattern frequencies, and hence high velocities, is expected to help maintain the stability of the optomotor feedback loop; at high gains a purely proportional controller would produce divergent oscillations [Warzecha and Egelhaaf, 1996].

Insect optomotor responses do not demonstrate complete compensation for drift of the surrounding panorama. The relationship between the open-loop and closed-loop responses at equilibrium can be calculated with reference to the feedback diagram (see also Section 7.4).

2.5.2 Other Motion Responses

Visually responsive neurons recorded in the neck connectives of *Acheta domesticus* were recognised early on as being physically and functionally similar to movement detector neurons in the locust [Palka, 1972]. In the locust the lobula giant movement detector (LGMD) and its post-synaptic target, the descending contralateral movement detector (DCMD) are tuned to mediate fast escape manoeuvres from looming objects [Rind and Simmons, 1999]. Whilst the cricket neurons have not yet been tested with looming stimuli they display the same characteristic non-directionally-specific response to motion of small objects, which is suppressed during wide-field motion. This suppression is peripherally-generated [Rowell et al., 1977] and therefore occurs in response to both ex-afferent and re-afferent optical flow. This mechanism ensures that the cell remains unresponsive during self-generated rotations, illustrating one type of solution to the reaafference problem. By contrast, the system mediating optomotor following would clearly require a different type of mechanism, since wide-field optical flow is exactly the stimulus that the optomotor system needs to respond to.
2.5.3 Object Responses

The most thoroughly investigated visual response in crickets has been the fixation and approach of dark objects (scototaxis). The behavioural relevance of the response is not obvious, since dark coloured objects could correspond to various types of environmental features. Wood crickets are reported to walk towards the darkest area of their visual field in order to move from open fields towards their preferred habitat of forest borders [Honegger and Campan, 1989]. Dark areas of the visual field may also correspond to shelter from predators; Gryllus integer females offered a choice between two sources of calling song were most likely to turn towards the one approached through a covered area, even if the calling song in this direction would ordinarily be the less attractive [Hedrick and Dill, 1993]. Various attempts have been made to determine innate preferences for certain shapes and dimensions of targets and to relate these to habitat or life-style. For example Acheta domesticus has been found to prefer horizontally orientated stripes [Atkins et al., 1987] which might correspond better than vertical shapes to potential shelters. There have been some conflicting findings, and there is no reason why the crickets' preference for certain shapes should not change over time. However the key finding, that dark targets can be attractive, is not disputed.

The mechanism of detection and fixation of dark objects during scototaxis is not known. Crickets walking towards dark targets in an arena setting keep an edge of the target in a certain region of the visual field by rotating the body [Lambin, 1984]. Lambin suggests that the reason the target edge is allowed to move and not fixed more accurately during tracking may be to prevent adaptation of photoreceptors. Experiments on crickets that have been tethered with their heads fixed supports this view; orientation was abolished under constant light but regained if the light was flashed on and off [Goulet et al., 1981]. Freeing the head additionally allowed the crickets to perceive relative distance. Alternatively the model put forward by Wehrhahn and Hausen for tracking and fixation in Musca [Wehrhahn and Hausen, 1980] produces fixation as a by-product of motion responses and may apply to the cricket as well as to the fly.

2.6 Auditory & Visual Interaction

2.6.1 A Motivating Example

The kind of multimodal interaction considered in this study is illustrated by an example from [Weber et al., 1981]. Crickets were tested on a walking compensator for orienta-
tion towards standard calling song broadcast from one of two speakers, separated by an angle of 135°. Occasionally the active speaker was switched so that, assuming the cricket was still responding, it would re-orientate to the new source of the sound.

During one such trial the effect of switching the room lights on and off was tested. Switching the lights on revealed some of the structure of the anechoic chamber in which the experiments were being performed, including the speakers, some light bulbs and parts of the walking compensator apparatus itself. When tracking a speaker in the dark the crickets were seen to meander by up to 60° either side, but when the lights were switched on the crickets tracked more accurately, deviating by only 6°.

Knowing that crickets have an optomotor response, and given its likely rôle in course stabilisation, it seems reasonable to think that it could be responsible for the change. The introduction of visual contrast would have permitted optical flow cues to be extracted, and any turns away from the speaker direction would have been counter¬acted.

The problem with this idea is that the optomotor following response would equally well counteract the turns of the cricket towards the speaker. The example shown by [Weber et al., 1981] (their figure 5A) provides some evidence to suggest that this doesn’t happen. It includes two examples of the sound direction being switched, one in the light and one in the dark. When the switch happens in the light the cricket continues to run in the direction of the initial speaker for around 15 seconds, but then re-orientates itself to the new speaker within 5 seconds. When the same kind of switch is made in the dark the cricket starts to re-orientate much sooner, although the speed of turning is not much different.

The central question addressed by the experiments in the later chapters of this dissertation is: What features or mechanisms in the brain permit the optomotor reflex to function without impeding turns to sound? No published experiments have sought to answer exactly this question for crickets. A number of possibilities are suggested by the results of previous behavioural experiments, and these are outlined below.

### 2.6.2 Summation of Reflexes at the Motor Output

When it is stated above that the optomotor system would counteract the turns of the cricket towards the speaker it is assumed that the pathways mediating phonotaxis and optomotor following function in parallel, before their effects are summed at the motor output. This is illustrated in the block diagram of Figure 2.9. The flow lines in the
Chapter 2. The Components of Insect Behaviour

diagram can be considered to carry one-dimensional steering signals, where positive values represent anticlockwise turning, negative values clockwise turning, and zero represents heading straight. The functions computed in the boxes are assumed not to reverse the sign.

![Diagram](image)

Figure 2.9: Parallel operation of the phonotaxis and optomotor systems with summation at the motor output.

By following the signal flow from the auditory system through the diagram it is clear that any turn to sound immediately causes visual feedback and an opposing signal on the forward optomotor pathway.

Even so, it is still possible that this organisation could exist in the cricket brain without being detrimental to phonotaxis. In the most trivial situation the optomotor system might just produce an output so small that its effect on phonotaxis could be neglected; indeed for visual situations where there is no scene structure this must be the case. In the trackball experiments of Chapter 4 data is collected to measure the magnitude of optomotor turning and confirm that this is not the case in high-contrast visual environments where crickets can, nonetheless, orientate to sound very successfully.

A more complex possibility is that the frequency characteristics of phonotactic steering and the optomotor response are such that the two systems can function in parallel, as suggested for the fly by [Egelhaaf, 1987]. Male flies find and follow mates in visually guided aerial chases. During these chases the pursuing fly makes zig-zag turns. As for the cricket, it had been supposed that the optomotor response would interfere with these turns. However, a study of the filtering properties of the optomotor system showed that it acted as a low-pass filter. The high-frequency pursuit turns occurred at a rate above its cut-off frequency. Egelhaaf proposed that this would avoid the two systems conflicting, although it has also been pointed out that short discrete pattern movements still contain low frequency components (S. Laughlin, personal dis-
The approach taken in the current study to investigating this possibility is to incorporate temporally scaled models of cricket phonotaxis and optomotor following onto the robot of Chapter 6, then check how they function together in closed-loop tests of orientation to sound.

2.6.3 Chained Subsystems

The idea of the chained sub-systems model is that two sensorimotor systems are connected in series, so that the second one has full control over the motor output. This seems to be an arrangement that exists in part of the fly brain. In flies the halteres are an essential input for flight control, acting like gyroscopes to detect accelerational forces and transmit correctional reflexes to the wings. Unintuitively, they may also play a direct role in the visual control of flight. Recordings from haltere control muscles show clear responses to visual motion stimuli [Chan et al., 1998]. This suggests that certain descending visual neurons may effect turns during flight by modifying the halteres’ directional sensitivity. The haltere’s sensory fields would respond as if to an accelerational change, relaying the visual command on to the wings. This hypothesis is corroborated by the absence of visual motion responses in the wing control muscles during quiescence.

![Diagram](image)

Figure 2.10: "Chaining" of the phonotaxis and optomotor systems by injection of the auditory signal directly into the optomotor system.

An analogous arrangement for the cricket behaviours considered here would be for the output of the phonotaxis system to be injected into the visual system as if it were an optical flow input. Clearly this could not happen immediately at the sensing level, but it is assumed that the auditory pathway would join the optomotor pathway early enough that its signals would undergo most of the filtering that takes place for visual
signals. This is illustrated in Figure 2.10. Note that the visual system mediates the motor control here, whereas it takes the opposite rôle in the fly haltere example.

This new “chained subsystems” model is still closely related to the parallel scheme of Figure 2.9, but the transmission of auditory steering signals via the visual system implies that the optomotor filter must be capable of passing all the frequency components present in the auditory steering signal. This is tested in Section 4.5.2.

2.6.4 Inhibition

Another straightforward modification of the summation scheme of Figure 2.9 is to imagine that the cricket follows the rule “when turning to sound, ignore the optomotor system”. This is illustrated in Figure 2.11, where a subsidiary pathway, branching from the auditory pathway, is used to gate the optomotor system’s output. Evidence exists to show that houseflies use a similar approach to counteract their optomotor response when pursuing small targets over a patterned background [Srinivasan and Bernard, 1977].

![Figure 2.11: Complete inhibition of the optomotor pathway by the phonotaxis pathway.](image)

2.6.4.1 Outright Inhibition

If one were to try and implement the scheme of Figure 2.11, it would become apparent that the figure is ambiguous, since it is not obvious how the switch element should behave. The most crude approach would be to suppose that the branching pathway does not carry directional information, but just a signal that operates the switch whenever calling song is detected. This would have the desired effect of stopping the optomotor response from interfering with turns to sound, but efferent information would also be ignored. The cricket would even lose the benefit of the optomotor system’s stabilising effect when the sound was straight ahead.
This possibility is not incompatible with the example cited in Section 2.6.1, since the stabilising effect seen there was not necessarily due to optical flow. A certain type of bush cricket, *Poecilimon affinis*, also performs more direct phonotaxis on a walking compensator in the presence of visual structure [Von Helversen and Wendler, 2000]. Helversen and Wendler found that a fixed landmark at any position on the horizon allowed *P. affinis* specimens to maintain their heading towards the speaker during periods of silence. If such visually-guided tracking was established, then the landmark was suddenly displaced with a discrete jump, the bushcrickets would re-align themselves on the same relative bearing. This suggests that they were using a memory of their static visual surroundings to stabilise themselves, rather than motion cues. It is not known whether *Gryllus* use such a strategy in similar circumstances.

### 2.6.4.2 Directionally-Sensitive Inhibition

A more sophisticated kind of inhibition might make use of the directional information in the auditory pathway. If the inhibition were to operate only when the directional signal indicates that the sound source is located laterally, then the more useful effect of ignoring optomotor commands only when turning to sound would be achieved. In practice a cricket might face the sound exactly very infrequently, and the spatial resolution of the auditory system might not be precise enough to detect it, so a more practical rule would operate the switch when the absolute value of the steering signal is above some threshold.

A further refinement to the rule can be envisaged, on the basis that the optomotor system normally operates under conditions of negative feedback; anticlockwise turning results in clockwise optical flow and *vice versa*. If anticlockwise turning results in anticlockwise optical flow then there is probably an important external disturbance, which should be responded to. So a rule might exist to ignore optical flow only in the "expected" direction. All these kinds of rules could in theory be distinguished by quite simple experiments with open-loop multimodal stimuli as discussed in Section 2.7.2.

### 2.6.4.3 Interleaved Inhibition

So far it has been assumed that the inhibitory signal is continuous, however the cricket calling song has a chirp and syllable structure that can persist all the way through to the steering movements of the animal [Hedwig and Poulet, 2004]. If the hypothetical inhibitory signal were to retain this structure then this would be to the cricket’s advan-
tage, since the inhibition could take place at exactly those times when turning takes place. The optomotor response would be able to perform its stabilising function in between chirps or syllables. Some information about continuous external disturbances would be ignored, but only for a proportion of the time.

Such interleaved inhibition might be more difficult to detect than the outright inhibition discussed in the previous sections, since in a long-term average the effect on a cricket's turning tendency would be smaller; in the standard synthetic chirp pattern used for the trackball experiments of Chapter 4 the duration of the sound, and hence the proposed inhibition, is only 25% of the chirp repetition interval. One of the contributions of the experiment described in Section 4.4.2 is to show that optical flow has an effect even when it is restricted to exactly those times when the cricket is turning. This provides some reassurance that the time-averaged tests of Section 4.4.2, which were initially conducted to look for an inhibitory effect, did not fail to detect one just because of its short duration.

2.6.5 Modulation of the Optomotor Gain

An experiment described in [Böhm et al., 1991] examined the effect of combining phonotaxis and scototaxis in *Gryllus bimaculatus* on an open-loop trackball. A characteristic curve for phonotaxis was produced by moving the speaker in discrete steps around the cricket. The same technique was then repeated with an attractive black target as the stimulus. The curve for turning towards the speaker was sinusoidal, but the curve for turning to the stripe was like a modified sinusoid, with the slope of the curve steeper in the frontal region. When the calling song and the stripe were moved around the cricket together, the resulting curve was best fitted by summing the phonotaxis/scototaxis curves with the ratio 1:0.7.

One interpretation of this result is that the cricket made a choice to pay more attention to the sound, by reducing the weighting of its visual turning signal. This implies something other than an outright inhibition, that might be better described as modulation of the gain of the visual pathway. This type of mechanism, applied to the optomotor pathway, is pictured in Figure 2.12. It should be noted that Böhm and colleagues’ experiment would not have distinguished between gain modulation and interleaved inhibition as described in the previous section; over the course of several chirps interleaved inhibition would also have resulted in a partial reduction of the visual component. The difference would lie in the neural mechanism that underlies the
interaction between the two pathways.

Figure 2.12: Adjustment of the gain of the optomotor pathway by the phonotaxis pathway. $k$ — gain of optomotor system.

There is another reason, however, for treating gain modulation separately from inhibition: a gain modulation mechanism might act under some circumstances to increase the efficacy of the optomotor pathway. This seems to happen in the case of a bush cricket, *Requena verticalis*, as shown by the experiment of [Bailey et al., 2003]. Bailey and colleagues tested *R. verticalis* for orientation in a darkened arena. Ordinarily the bushcrickets were not attracted to a small lightbulb. However, if the lightbulb was placed over an active speaker which was broadcasting calling song, then the bush crickets would follow the lamp around once the sound was stopped. Unlike *Poecilimon affinis*, described in Section 2.6.4.1, *R. verticalis* would not follow the lightbulb if it made a discrete jump; the movement had to be smooth. This suggests that they were relying on motion cues rather than a memory of their visual surroundings. It appears that the initiation of phonotaxis had the effect of increasing the gain of an optomotor-like system to help the insects “lock on” to the direction of the sound.

In the type of multimodal behavioural experiments discussed in Section 2.7.2 the effect of gain modulation would be largely the same as for inhibitory schemes. In practice the tests of Section 4.4.1 did not reveal any evidence of this type of mechanism, so it was not necessary to try and distinguish between them further.

### 2.6.6 Efference Copy

The concept of an efference copy was proposed as a solution to exactly the problem under consideration: that of how an animal can make intentional turns without its optomotor response interfering [Holst, 1973]. Holst reversed the optomotor feedback loop of flies, which displayed continuous circling behaviour as a result. This indicated that
the flies continued responding to their visual input during turning, resulting in positive feedback. Since the flies were apparently not inhibiting their optomotor system Holst was led to propose an alternative mechanism by which they could avoid being immobilised by their optomotor responses under ordinary conditions. The general idea is that the motor pathway responsible for initiating non-optomotor turning branches, sending a copy of the efferent motor command (hence “efference copy”) to the sensory region of the brain where it interacts with the afferent sensory signal, cancelling out that portion due to re-afference.

Figure 2.13 shows how the efference copy principle might work in the cricket. The components of the parallel scheme of Figure 2.9 are still present. The difference is that the phonotaxis pathway sends off a second branch before the original point of summation, which is modified by a box labelled “predictive model”, before being added to the output of the optomotor system.

![Figure 2.13: Efference copy from the phonotaxis system to the optomotor subsystem. Dashed line indicates feedback error signal, which may modify the predictive model if it is adaptive.](image_url)

**Efference Copies vs. Corollary Discharges**

In layout the efference copy scheme is not so different from the inhibitory schemes considered in Section 2.6.4. The difference lies in the fact that the efference copy is transformed into a signal that predicts the size and direction of the re-afferent component. A more general concept that might be taken to encompass both efference copy and some kinds of inhibitory schemes is corollary discharge [Crapse and Sommer, 2008]. According to Crapse and Sommer:

> The term “efference copy” implies an actual copy of the motor command (the efference) that targets the muscles […]
whereas

In studies on fish, Sperry coined the term “corollary discharge” to denote motor-related signals that influence sensory processing, but his conception was less specific as to where the branch from the motor to sensory pathways should emerge […] At a mechanistic level corollary discharge can adopt one of multiple forms depending on how it is used: it can facilitate, inhibit or otherwise modulate its target.

So all three schemes discussed in Section 2.6.4, gain modulation and efference copy, would fall under this definition. In this dissertation the term efference copy is taken to mean a corollary discharge signal that is predictive of the magnitude of the reafference.

Forward Models

The box labelled “predictive model” in Figure 2.13 is a potentially important component, although its contents could vary widely in complexity. In Holst’s original formulation of efference copy this transformation was not explicitly considered. However in most situations a straightforward copy of the motor signal is unlikely to be a good predictor of the sensory re-afference. Temporarily assuming an absence of exafferent optomotor stimuli, the output of the optomotor system is a product of the history of the auditory system’s output, the dynamics of the muscles, and the visual properties of the environment. The result is that the re-afference is likely to lag the efference by some interval, have a different frequency spectrum, amplitude and co-ordinate system. The function of the predictive model is to transform the efference copy to take these factors into account.

That said, from an evolutionary perspective a predictive signal only has to be good enough to confer selection advantage to the insect. The time course of an efference copy may in some cases be a close enough match to the sensory re-afference that the necessary transformation can be accomplished using very little neural hardware. This is illustrated by an identified multisegmental neuron in the ventral nerve cord of the cricket [Poulet and Hedwig, 2006]. This neuron is driven by the motor circuitry that produces the calling song in males. The neuron spikes in time with the sound pulses produced by the singing network, and directly inhibits both primary auditory afferents and auditory interneurons in the prothoracic ganglion, preventing desensitisation of the auditory system during singing. In this case the properties of a single neuron are sufficient to produce inhibitory signals that have an appropriate timing and amplitude.
Adaptation

An example of the potential complexity of the forward modelling component in efference copy comes from weakly electric fish [Bell, 2001]. Mormyrid fish have electroreceptors in their skin that detect external voltage sources (passive sensing). They can also produce electric organ discharges (EODs), which cause pulses of current to be emitted. Impedance of this current by nearby objects in the water modifies the re-afferent signal, indicating the presence of possible prey (active sensing). The two types of sensing are ultimately mediated by two specialised types of receptors, but the ampullary receptors that mediate passive sensing also respond to the EOD. In order to filter out the effects of the self-generated current the ampullary afferents synapse onto the principle cells of a cerebellum-like structure in the electrosensory lobe. There, an efference copy associated with the EOD motor command is transformed into a negative copy of the sensory reaference. This can be observed by blocking the EOD with curare and recording from the principal cells.

The predictive signal is adaptive. By repeatedly pairing the corollary discharge with an artificial electrical stimulus at the skin it is possible to observe the prediction changing to match. The cerebellar-like structure performs the function of detecting and adapting to sensory signals correlated with the efference copy. The prediction matches the lag, temporal pattern, magnitude, and spatial position of the re-afference.

It is not known whether insects use forward models of such complexity. The example of the corollary discharge interneuron described in the previous section appears to be “hard-wired”. Under ordinary circumstances the cricket calling song is so stereotyped, and the feedback loop from song pattern generator to ears is so tight, that there would be no need for adaptation. In the mormyrid fish adaptation can compensate for changes in the reafference due to water conductivity or other aspects of the fish’s immediate environment. For the cricket’s optomotor response the situation might be similar. As the cricket moves around its environment different visual surroundings will produce differing levels of motion cue. Therefore adaptation of an optomotor efference copy might be necessary.

Detecting an Efference Copy

In the example of the electric fish discussed in the previous section, the presence of an efference copy was shown using electrophysiology. The signal was recorded directly at the point where it converged with the incoming pathway from the electrosensory
cells, whilst the production of the reaferent signal that would normally cancel it out was prevented. In the case of the cricket there was no firm evidence to suggest that effort should be put into searching for a neural correlate of an efference copy. Rather, it was a behavioural indication that such a mechanism might exist that was sought. How might this be found?

A cricket walking on the open loop trackball is actually in a situation closely analogous to that of the fish in the efference copy experiment. If the cricket is phonotactically responsive then a chirp of calling song will cause an auditory steering reflex and, assuming the efference copy model is correct, a corollary discharge. This will converge with the optomotor pathway, and, since there is no optomotor input, appear (summed with the auditory signal) in the turning behaviour.

Assume that the cricket’s hypothetical predictive model transforms a corollary discharge indicating rightward turning into a prediction of leftward optical flow and vice versa. Then by following the signal pathways of Figure 2.13 it can be seen that the effect of the absence of the correct optomotor feedback would be to make the cricket turn more strongly in the direction indicated by the auditory system. The problem is that this gross change in behaviour is exactly the same as that predicted for summation at the motor output or the chained sub-systems hypothesis. There too, the removal of the expected negative optomotor feedback would result in stronger turning to the sound.

If the predictive model in the efference copy mechanism is adaptive, then a more powerful possibility emerges: cause the predictive model to adapt into a new state, and the cricket’s response in a previous state should change. This possibility is exploited in the tests of Section 4.4.2.

### 2.6.7 Summary

Five different categories of mechanism that a cricket could use to overcome optomotor reafference during phonotaxis have been illustrated. These are:

- Summation at the motor output (Figure 2.9)
- Chained sub-systems (Figure 2.10)
- Inhibition (Figure 2.11)
- Modulation of the optomotor gain (Figure 2.12)
• Efference copy (Figure 2.13)

These five schemes are taken as alternate hypotheses for investigation throughout the rest of the dissertation.

Only one published experiment directly addresses auditory and optomotor combination in *Gryllus* [Böhm et al., 1991], using an open-loop trackball set-up. Whilst the results of that study were ambiguous the experiment was a reasonable one in principle. The following section reviews the data it produced and explains, using the results of a published modelling study [Webb and Reeve, 2003], the extent to which replicating the test could help distinguish between the five hypotheses stated above.

### 2.7 Behavioural Clues to the Mechanism of Interaction

#### 2.7.1 Open-Loop Tests [Böhm et al., 1991]

Böhm, Schildberger and Huber examined the effect of combined auditory and optomotor stimuli on *Gryllus bimaculatus* using an open-loop trackball. Characteristic curves for turning to calling song at different angles to the cricket’s midline were calculated in the presence of a stationary grating as well as with the grating moving either left or right. In a second experiment a characteristic curve for optomotor following was calculated for pattern frequencies up to 5 Hz, with a calling song being played from 0° or 60° either side of the cricket’s midline. The recorded turning tendency for each stimulus combination was a mean calculated over two minutes. The resulting curves are shown in Figure 2.14. The authors state:

[... ] any combination of optomotor and sound stimuli caused a turning tendency that can be explained as the weighted sum of the turning tendencies evoked by the two individual stimuli.

There are two points that should be made regarding this experiment. Firstly, although the authors claim to see overall turning tendencies that are a weighted sum of those evoked by the individual stimuli, they don’t provide any data to show the effect of optomotor stimuli alone. This is important to know since, as discussed in Section 2.6.4, it could be that the presence of calling song influences the optomotor reflex independent of the sound direction.

In the left-hand graph the central line shows the characteristic curve for turning to sound in the presence of a stationary grating. Each of the outer lines shows the
2.7. Behavioural Clues to the Mechanism of Interaction

Figure 2.14: Left — Auditory characteristic curve in the presence of a stationary grating, a grating moving right with a pattern frequency of 1 Hz, and a grating moving left at 0.5 Hz. Right — Optomotor characteristic curve for pattern frequencies between zero and five Hertz, with a calling song at an angle of 0° or 60° either side of the midline. Reproduced with permission, from [Böhm et al., 1991].

effect when, in addition, the grating is rotated in one of two directions (note that the speeds of rotation are not identical). Since the effect of a constant optomotor stimulus is expected to be constant on average the “weighted sum” claim implies that the curves should show a parallel shift.

The second point to note, as highlighted in [Webb and Reeve, 2003], is that the curves are not truly parallel. Each point of the graph was collected from three crickets, so inter-individual variability could account for this to some extent. However it is notable that, for instance, the effect of the leftward grating movement when the sound is straight ahead is a shift of around 18°s. By contrast, when the sound is to one side the shift induced by moving the grating seems to be reduced in many regions to around 10°s or less. This kind of difference could indicate the presence of an inhibitory mechanism that acts when the cricket is turning, of the type discussed in Section 2.6.4.2.

Since no cue to the variability of each point is given on the graph it is not possible to say whether such observations are significant. [Webb and Reeve, 2003] showed that if data were available to test the significance of such shifts, then it might be possible to distinguish the presence of an inhibitory mechanism from the other schemes, as will now be discussed.
2.7.2 The Effect of the Five Mechanisms in Open Loop Tests [Webb and Reeve, 2003]

Webb and Reeve used a simple numerical simulation to examine the expected effect of different mechanisms of auditory / visual combination on the results of Böhm, Schildberger and Huber’s experiment. The mechanisms studied were largely the same as those outlined in Section 2.6: summation at the motor output, chained sub-systems, inhibition and efference copy. Modulation of the optomotor gain was not considered. Rather than calculate a full characteristic curve for phonotaxis only three speaker positions were considered: 0°, and ±60°. The sets of curves produced for the different simulated controllers are shown in Figure 2.15.

![Figure 2.15: Simulated open-loop multimodal experiments. Reproduced with permission, from [Webb and Reeve, 2003].](image-url)
2.7. Behavioural Clues to the Mechanism of Interaction

No Opto

The top-left axes show the effect of the auditory system alone, when the optomotor system is given zero gain. This shows that the behaviour of the simulated cricket is to turn anticlockwise for speaker positions to the left, clockwise for speaker positions to the right, and to not turn at all when the speaker is at 0°.

Summation at the Motor Output

The top-right axes show the behaviour under summation at the motor output (termed “addition” in the original paper). With reference to the block diagram of Figure 2.9 it can be appreciated how, with the visual feedback loop broken, a constant optomotor input of positive sign will cause an equal positive shift for each point of the “no opto” graph, and a constant negative optomotor input will cause an equal negative shift.

Chained Subsystems

The chained subsystems case (termed “follow-on” in [Webb and Reeve, 2003]) is illustrated by the bottom-right graph. The simulation shows that the effect in time-averaged open-loop conditions is for a constant optomotor stimulus to cause a parallel shift in the auditory-only curve, indistinguishable from the effect under summation at the motor output. To understand the reason for the parallel shift, consider the block diagram of Figure 2.10. Denote the signal entering the phonotaxis pathway by \( p \) and assume it is subject to filtering by a filter \( F_p \), before it is injected into the optomotor system. Assume the optomotor system has its own filter \( F_o \), which receives the output of the phonotaxis system added to the optomotor input \( o \). Now the simulated cricket’s turning tendency, ignoring the thoracic motor network and legs, is given by \( \phi = F_o(F_p(p) + o) \). In the simulation \( F_o \) is a linear low pass filter (which is a reasonable choice in view of the experiment of Section 4.3.3), so by additivity of linear filters \( \phi = F_o(F_p(p)) + F_o(o) \). The optomotor contribution is independent of the phonotaxis component, and is constant for constant \( o \).

Inhibition

Both the central graphs in Figure 2.15 show variants of an inhibitory controller. As discussed above, the optomotor pathway contained a low-pass filter, and the two versions differ in whether the inhibition was applied at the input (pre-integration) or the output (post-integration) of that filter. As for Figure 2.11 the inhibition was absolute
when it was applied. However, the inhibitory signal matched the chirp structure of the calling song, so it was only active for part of the time. Furthermore it was directionally sensitive in the first sense described in Section 2.6.4.2; it operated only when the speaker direction was outside the frontal field, without distinguishing between left and right.

For the particular (linear) optomotor filter used in the simulation the effects of the two inhibitory schemes were quantitatively identical. When the sound was at 0° the inhibitory signal did not act, and the effect of the optomotor inputs was the same as in the additive case. When the sound was at ±60° the inhibitory signal was active during each chirp, and so short periods of turning to sound only were interspersed with periods when the turning tendency included a component due to the optomotor pathway.

If the inhibition was applied post-integration the output signal of the optomotor filter reached a steady state, and this signal was gated, giving a series of square pulses. If the inhibition was applied pre-integration the input to the optomotor filter was a series of square pulses, which resulted in an output that settled exponentially towards each new input level. When either type of output pattern was averaged over a few cycles of the chirp pattern the effect was the same: an optomotor contribution of lower magnitude than in the additive case. When plotted on the graphs of Figure 2.15 this corresponds to a reduced shift between the central curve and the two outer curves.

It is not the case that in general the pre- and post-integration inhibition schemes should produce exactly the same output. Qualitatively, however, they will always produce a reduction in the size of the shift induced by an optomotor stimulus.

**Efference Copy**

The bottom-left graph of Figure 2.15 shows the output of the simulation using the efference copy scheme. It should be noted that the efference copy mechanism used in [Webb and Reeve, 2003] was different from that shown in Figure 2.13: it did not use any predictive model, only a direct copy of the motor command, and the branched pathway was combined with the optomotor pathway before the optomotor filter. This is illustrated in Figure 2.16.

Using a similar argument to the chained subsystems case, denote the phonotaxis filter as $F_p$ and the optomotor filter as $F_o$, and the auditory and optomotor input signals as $p$ and $o$ respectively. It can be seen that the output of the system (assuming the feedback loop is broken, the optomotor filter is linear and ignoring the thoracic motor networks and legs) is $\dot{\phi} = F_p(p) + F_o(F_p(p)) + F_o(o)$. Once again the optomotor com-
Figure 2.16: Efference copy from phonotaxis system to optomotor system as described in [Webb and Reeve, 2003]

ponent is independent of the phonotaxis input, and constant. This results in a parallel shift of the curve when the optomotor input is altered. The argument is not altered fundamentally for the arrangement shown in Figure 2.13.

2.7.3 Summary

Webb and Reeve’s simulations showed that Böhm, Schildberger and Huber’s original test might have revealed the presence of a directionally-sensitive inhibition mechanism, had their data not been so affected by inter-individual variability. Summation at the motor output, chained subsystems or efference copy are all expected to produce parallel shifts in the curves. It is even possible that a non-directionally-sensitive inhibition mechanism could produce a parallel shift in the open-loop experiment. However, that shift would be less than the response induced by grating movement alone, which can be tested for.

Modulation of the optomotor gain was not considered in [Webb and Reeve, 2003], but the same logic applies as for inhibition: if the gain modulation is directionally-sensitive then this will result in a non-parallel shift of the curves, and if not, then the size of the shift will be different (possibly larger) than that induced by moving the grating in the absence of sound.

2.8 Detailed Modelling

Phonotaxis and optomotor following as individual behaviours have both been the subject of investigation through more detailed modelling, including robotic modelling.
Since the goal of the current project was to study the interaction of the two behaviours the approach taken was to assess the suitability of the existing models and make any alterations necessary to achieve a satisfactory match to the cricket.

Overall, insect biorobotics has tended to focus on problems relating to single sensory systems. There have been a few exceptions, for instance the Sahabot [Lambrinos et al., 2000] was used to investigate the capabilities of different types of visual sensor for navigation tasks. A panoramic camera and polarised light sensors were built onto the same wheeled robot and were used to perform homing in the natural environment of the desert ant Cataglyphis. However, these systems seem to represent complementary methods for navigation, and were not really integrated in any functional sense.

Section 2.8.3 focuses on three of the most relevant studies that have sought to integrate optomotor following with other behaviour. The results of these projects are reviewed, and the areas in which their biological relevance could be improved are highlighted in Sections 2.8.4 and 2.8.5.

### 2.8.1 Phonotaxis

A series of mobile robots that model phonotaxis without consideration for multimodal aspects have been developed. The first version tracked calling songs with the carrier frequency and syllable rate adjusted to take into account the limitations of the auditory processing hardware [Webb, 1993]. Its controller located the sound by integrating pulses from the two ears in leaky integrators and turning to the side of whichever integrator first reached a given threshold. The integrators also provided a degree of SRI specificity, since they were required to drop back below threshold before they could trigger another turn.

Subsequent models used a Khepera miniature robot as the platform, together with a custom-built ears circuit capable of responding to natural cricket calling songs [Lund et al., 1998]. Using this robot new controllers were implemented and tested, using progressively more realistic neuron models, and more detail based upon the identified circuitry of the cricket brain. These controllers provided a better match to the band-pass response to syllable rate [Lund et al., 1998], replication of behavioural experiments with song from above (cricket experiments described in [Stabel et al., 1989]) and a demonstration of selection of the better song in choice tests [Webb and Scutt, 2000]. The motor control in these models was relatively basic, with
three modes (no turn, left turn and right turn) in which turns were of a fixed duration (400 ms).

The most recently published model [Reeve and Webb, 2003] includes details of the connectivity of the auditory neurons in the prothoracic ganglion as well as the band-pass filtering properties of neurons in the brain (see Section 3.2.2). This robot employed a more complex mode of motor control in which the speeds of the wheels on each side of the robot were set independently. A pair of recurrently connected neurons provided bursts of activation that drove the robot forward with a bout structure. Spikes in the left or right brain neurons contributed excitation to the respective motor neurons, but there was no discrete switch between left and right turning. Instead the turning speed was determined by the relative level of excitation in the two sides.

An alternative approach to modelling phonotaxis has been applied to the investigation of choice paradigms [Mhatre and Balakrishnan, 2007]. *Plebeio Gryllus guttiventris* were videoed in an arena, performing phonotaxis in the presence of two aphasic conspecific calling songs coming from different directions. From the video recordings the stops made by each cricket were identified. Then, for each stop, the volumes at the cricket’s position of the sounds from each speaker were calculated. These were modified according to models of the cricket’s peripheral auditory directionality, and assessed for masking effects [Pollack, 1988]. Statistical descriptions were built up of the pause and bout durations performed by each cricket, as well as turn sizes and directions based on the stimulus situation. These were then used as the basis of a simulation in which the model cricket moved using a cyclic program: pause, measure song SPL (sound pressure level), walk. The pause durations, bout durations and turn angles were all sampled randomly from the pre-determined distributions, parameterised by the sound field properties where appropriate. The resulting paths were a good match for natural tracks of *P. guttiventris* on measures of direction, angular deviation and sinuosity.

**2.8.2 Insect Visual Behaviour**

Consistent with the relatively small amount of biological research into cricket vision there have not been any modelling studies that have addressed visually-mediated behaviour in the cricket specifically. However, similar behaviours in other insects have been given rather more consideration. Optomotor following is by far the most important of these in the current context. The maturity of the theoretical work relating to the
The Hassenstein-Reichardt (H-R) elementary motion detector (EMD) model has meant that a number of robotic implementations have been built. The studies discussed below are intended to show what mechanisms could be adapted for use in an optomotor response for the cricket robot, and therefore focus upon a rather limited subset of the available literature.

A number of other insect visual behaviours have been built into robots, including depth perception [Lewis and Nelson, 1998], navigation using centring responses based on bees [Srinivasan et al., 1999a], and collision avoidance using models of the type of neurons discussed in Section 2.5.2 [Rind, 2002]. It is likely that some or all of these apply to the cricket in relevant behavioural situations, but they were not required for the model under construction.

The Hassenstein-Reichardt Model

When a textured pattern moves in front of the compound eye of an insect the sequence of light intensities sampled by two photoreceptors lying along the direction of movement is correlated in time. This idea has been formalised in a mathematical model of movement detection that has proven to be a good description of the phenomenon of optomotor following in insects and other animals [Hassenstein and Reichardt, 1956] [Srinivasan et al., 1999b].

Figure 2.17 shows a schematic of the model. Two spatially separated intensity sensors correspond to ommatidia of the eye. The signal from each ommatidium is multiplied with a delayed signal from the other. The results of the two multiplications are subtracted, resulting in a directional signal. The output of such a detector in response to a moving grating shows various features that correspond to the behaviour of insect tangential cells, and the walking behaviour of insects on a trackball, including a log-normal dependency on the temporal frequency of the pattern [Srinivasan et al., 1999b]. The output is therefore not a reliable indicator of pattern velocity, and depends on the spatial frequencies present in a scene.

This motion detector model has been the subject of a great deal of theoretical analysis which it is beyond the scope of this dissertation to cover. There is no comprehensive review, but see [Egelhaaf et al., 1989] for a mathematical analysis of the standard model in response to moving sinusoids and references to older work.
Figure 2.17: Outline of the functioning of an elementary motion detector presented with a pattern moving from left to right. The signal received at receptor 2 is delayed compared to that received at 1. The signals from the two receptors are both delayed by equal amounts (τ), resulting in signals 1' and 2'. The speed of pattern movement meant that the lag between 1 and 2 roughly equalled τ, and the signals at 2 and 1' are strongly correlated. When multiplied (M) this results in a large positive signal. 1 and 2' are not strongly correlated, and multiplication produces only a small signal. The results of right-hand multiplication is subtracted from the left, giving a positive output. This final subtraction means that pattern movement in the opposite direction would result in a negative output signal. Reproduced with kind permission from Springer Science+Business Media: Journal of Comparative Physiology A, Neural Networks in the Cockpit of the Fly, Vol. 188(6), 2002, Page 423, A Borst and J Haag, Figure 6A
An aVLSI Motion Detector Array

The most straightforward way of providing a robot with an optomotor response is to use a horizontal array of H-R motion detectors, the outputs of which are summed together to give a turning command. The signals from individual detectors display pattern-dependent modulation (e.g. a transient response to a passing stripe) but these are averaged out in the summation. For modelling purposes the output of the detector array is usually transformed by a low-pass filter to more closely approximate the neural or motor responses of a particular insect. For engineering applications requiring complete visual stabilisation a PI or PID controller could be employed instead.

Such an array was implemented in an aVLSI chip by Harrison and Koch, with the temporal parameters tuned to match those of Drosophila [Harrison and Koch, 2000]. The exact implementation of the motion detectors differed slightly from the scheme presented above. Inputs came from photodiodes coupled to photoreceptor circuits whose outputs were proportional to the logarithm of light intensity. Photoreceptor signals were then temporally filtered with a pass band between 2.8 and 10 Hertz. Cutting low frequencies modelled the effect of the lamina monopolar cells described in Section 3.3. These effectively ignore the ambient light level which contains no motion information. Cutting high frequencies reduced the effect of AC light flicker. The temporal delays (τ) were produced by the phase lag inherent in a low-pass filter, a common alternative to building a pure delay line.

The chip was fitted with a lens and tested on board a simple wheeled robot, where it was used to successfully compensate for a 5:1 motor asymmetry [Harrison and Koch, 1999]. It was also tested using a closed feedback loop to a video monitor, a set-up which had previously been used to present patterns to a real flies tethered to a torque meter [Warzecha and Egelhaaf, 1996]. Because the chip’s temporal parameters were tuned to match those of the Drosophila visual system a direct comparison could be made with the fly data. A notable feature of the chip’s output under both open and closed-loop conditions were large fluctuations, not correlated with the stimulus pattern, occurring at around 2 Hz. Oscillations at a similar rate are observed in Drosophila torque traces. The behaviour of the chip suggests that they could emerge from inherent properties of the motion detectors rather than intentional turning [Heisenberg and Wolf, 1988] or the properties of the optomotor feedback loop.

---

5Proportional-Integral or Proportional-Integral-Derivative
6Analogue Very Large Scale Integration
Extensions to the Basic Optomotor Model

A horizontal array of motion detectors across the whole or part of the visual field as described above can produce a good match to the optomotor following response. In flies, however, the motion inputs that control flight are not homogeneous over the visual field [Tammero et al., 2004]. The optomotor response still emerges from wide-field visual rotation because inputs from the eyes are bilaterally symmetrical, but when other patterns of optical flow are tested, different behaviours result.

Motion sensitive neurons in the lobula plate of flies respond with depolarisation to retrograde motion and hyperpolarisation to anterograde motion. The size of these effects are not symmetrical, with the former being larger [Egelhaaf et al., 1989]. It has been hypothesised that this could account for the tendency of flies to fixate a black vertical bar within their frontal visual field [Wehrhahn and Hausen, 1980]. Huber and Bülthoff used a robot very similar to the one in this study (see Chapter 6) equipped with a ring of EMDs. The bias in the directionality of the fly’s motion sensing was incorporated, along with a Gaussian noise process that generated continuous small rotations of the robot. Together these resulted in the fixation and approach of single high-contrast landmarks, while whole-field rotation produced the familiar optomotor following response [Huber and Bülthoff, 1998].

2.8.3 Integrating an Optomotor Response

Earlier Cricket Robots

Previous work on the problem addressed in this thesis — the integration of an optomotor response with phonotaxis — represents the most complete attempt at addressing a problem of multimodal integration in an insect biorobot. The first step towards this aim was an attempt to add the optomotor chip of [Harrison and Koch, 2000] to the phonotaxis robot of [Webb and Scutt, 2000] by adding the output of the two systems with suitable gains [Webb and Harrison, 2000]. This served to demonstrate that the two behaviours conflicted, and highlighted the need for an internal mechanism to co-ordinate the two systems. In order to produce a working system this additive scheme was adapted so that the optomotor gain was set to zero whenever the output of the phonotaxis system was non-zero. This modification enabled the robot to perform phonotaxis without turns being restricted. Under ordinary conditions the optomotor system did not improve the directedness of phonotaxis since there were no sources of
external disturbance for it to correct. However, when a bias was added to the motor system it produced significant improvements compared to the phonotaxis controller alone.

[Webb and Reeve, 2003]

A subsequent study [Webb and Reeve, 2003] took a more thorough approach, and considered alternative schemes that could be employed to integrate the two systems. The alternatives, and the initial simulation work were discussed in Section 2.7.2. Subsequently the algorithms used in the computer simulation were transferred to a Koala robot, where they were connected to the sensors used in [Webb and Harrison, 2000] to obtain more realistic inputs and verify the results of the simulation.

A final set of experiments used the more sophisticated neural model of [Reeve and Webb, 2003]. The optomotor chip’s analogue output was converted into spike trains that were fed directly into the neural model and used to excite a pair of mutually inhibitory optomotor neurons, one for each direction of motion. Summation at the motor output could be achieved by allowing the optomotor spike trains to excite the same motor neurons as the phonotaxis network. Chaining of the two subsystems worked by connecting the auditory brain neurons to the optomotor neurons instead of the motor neurons. Finally, inhibition worked by having each auditory brain neuron cross-connected with an inhibitory synapse to the opposite optomotor neuron. The brain neuron that triggered rightward turns inhibited the optomotor neuron that responded to leftward optical flow and vice versa.

Webb and Reeve note that the inhibitory neural network did not perform the same function as the algorithmic inhibition used previously, since the model synapses used would only suppress spiking in the optomotor neurons up to a certain level. Furthermore the inhibition was directional rather than outright, so the robot could still respond to visual motion in the unexpected direction during a phonotactic turn. In this way the neural version of inhibition had some features in common with the more complex efference copy scheme.

Tuning a Prediction Using a Genetic Algorithm

A modification of the inhibitory controller was described in [Russo et al., 2005]. In this version a genetic algorithm was used to tune the parameters of the synapses that transmitted spikes through the inhibitory pathways. The goal was to match the magni-
2.8. Detailed Modelling

titude of the resulting inhibitory potentials to the expected size of the reafference taken from neural recordings. This was successful in producing inhibition of the required magnitude and duration.

Combining Modelling and Biological Validation — [Collett, 1980]

The most clear limitation of the study of [Webb and Reeve, 2003] is the lack of biological data that could be used to validate the different models. A study by Collett used a combination of behavioural testing and simple cybernetic modelling to examine a similar type of interaction between smooth pursuit and optomotor following in the hoverfly Syritta pipiens [Collett, 1980]. Whilst the models were rather abstract by comparison to those of [Webb and Reeve, 2003], initially neglecting the dynamical properties of both systems, the study was ultimately more informative due to the better availability of biological information. Thus Collett was able to determine that the optomotor response was present during tracking, rejecting the inhibition hypothesis. It was also found that the gain of smooth pursuit stayed relatively constant over a broad frequency range, unlike that of optomotor following, leading to a rejection of the chained subsystems hypothesis.

Collett did go on to simulate the additive and efference copy models using analogue electronics. However the formulation of the efference copy scheme was like that of Figure 2.16; the efference copy signal was combined with the optomotor signal before any filtering occurred in the optomotor pathway. Furthermore, the contribution of the thoracic motor networks, effectors and environmental filtering were all omitted, making the sensory prediction task unnecessary. In the electronic model this meant that the efference copy was trivially a perfect match for the reafference. Collett claims that:

[... ] detailed modelling suggests that efference copy has several advantages over the additive scheme, and is certainly no more demanding in terms of complexity or neural hardware.

This is the case in the model, but only because the filtering effects of the sensory and motor systems were omitted. Collett suggests that this is reasonable in the case of Syritta, but the experiments described in Chapter 4 of this dissertation show that this is not a valid approach for Gryllus.
2.8.4 Improving the Match of the Dynamics

The dynamics of the robotic models of [Webb and Reeve, 2003] are much more complex, and are more relevant to the cricket than the abstract mechanisms of [Collett, 1980]. However they do have certain shortcomings for answering the question at hand. Firstly, the dynamics of the optomotor system are matched to Drasophila rather than the cricket. The plausibility of summation at the motor output and the chained sub-systems hypothesis depend critically on the frequency responses of the visual system compared to the auditory system. The internal behaviour of the phonotaxis model of [Reeve and Webb, 2003] is probably as accurate as it could be, given what is known about the neurons involved (Section 3.2.2). However the motor output is not modelled on any particular data, and could be improved, given what is known about rapid steering on the open-loop trackball (Section 2.4.3).

2.8.5 Implementing a True Predictive Mechanism

Of the different sensory integration schemes that were investigated in [Webb and Reeve, 2003] the efference copy scheme was never fully implemented on a robot. Even the variant that used an efference copy pathway tuned using a genetic algorithm [Russo et al., 2005] was not fully directional. In addition the biological plausibility of that version of the controller is questionable since the synapses used would have to have very long time constants in order to produce a sustained output. A different, biologically plausible, way of producing a full efference copy is explored in Chapter 6.

2.9 Summary

The literature reviewed in this chapter has shown that phonotaxis and optomotor following could be produced by relatively simple reflex-like pathways in the cricket nervous system. The most important features of sensory filtering and motor dynamics in these two systems have been highlighted from a functional perspective.

Five mechanisms that might underlie the integration of optomotor following with phonotaxis have been identified based on those investigated by a previous modelling study. These have been scrutinised alongside the available information on auditory / visual integration in crickets, and examples from the wider field of insect biology, with the following conclusions:
• The inhibition / gain modulation hypotheses might be distinguished from the others by replicating the experiment of [Böhm et al., 1991].

• The plausibility of the chained sub-systems hypothesis depends on the frequency responses of the auditory and optomotor systems. If auditory steering responses must pass through the optomotor system then it must not attenuate the relevant frequencies.

• In open-loop tests a non-adaptive efference copy would produce the same kind of effect as summation at the motor output or chained sub-systems. If the predictive component were adaptive then this might allow the efference copy mechanism to be distinguished without resorting to electrophysiology.

Previous modelling studies that have investigated similar hypotheses were reviewed, and the key areas in which their approaches could be improved upon highlighted:

• The dynamics of the motor systems in the previous robots of [Webb and Reeve, 2003] could be improved by incorporating the recent data on the fast components of phonotactic steering discussed in Section 2.4.3.

• Further data needed to be collected to correctly tune the optomotor input to the new cricket robot, rather than relying on a chip based on parameters gathered from Drosophila.

• The efference copy scheme had yet to be implemented fully using a biologically plausible mechanism.

• The study of [Collett, 1980] was made more conclusive by its combination of behavioural and modelling approaches, and this closer link with biological experimentation might have improved the more recent biorobotic investigations.
Chapter 3

The Physical and Neural Substrates of Cricket Behaviour

3.1 Introduction

Chapter 2 described phonotaxis and optomotor following using a functional approach that views the nervous system as composed of "black boxes" with certain inputs and outputs. This is a useful abstraction for trying to convey the different types of interaction that might occur between the two behaviours. When it comes to building a robotic implementation of the behaviours, however, a more detailed understanding is required.

For the robotic controller specified in Chapter 6 the approach taken was to make use of lower-level physical and neural detail in the model reflex pathways, whilst keeping the mechanisms of interaction between the two abstract and functional. This reflects the relative maturity of the current understanding about these two aspects of the insect brain.

In the individual reflex pathways the known physiology is relatively well established, and certain mechanisms are known to be essential to the behaviour. For instance the tympani of the cricket ear act as pressure difference receivers, which is essential for the angular acuity of the auditory system (Section 3.2.1). Motion detection in the insect optic lobe is almost certainly achieved by the correlation of signals sampled at neighbouring points of the retina (Section 3.3). The first two sections of this chapter put forward the uncontroversial findings that are used in the model to provide the best fit possible to the processes taking place in the cricket brain.

On the other hand the details of neurons and structures that are likely to be involved in multisensory co-ordination are far from clear. Recordings from neurons descend-
ing from the brain to the thorax are reviewed in Section 3.4. These have revealed certain general properties about the signals that control walking, but the identification of individual cells has proved difficult, and no details of connections to specific brain or thoracic neurons are known. This thesis speculates that the mushroom bodies of the insect brain could, amongst other less controversial functions, process corollary discharges to generate sensory predictions that modify the activity of descending neurons. In order to support this claim, the final part of this chapter reviews the known anatomy, physiology and behavioural functions of the mushroom bodies. These data, taken together, suggest that the elements necessary to process efference copies are present in the mushroom bodies. On the basis of this evidence the robot of Chapter 6 incorporates a predictive component based on a liquid state machine, an artificial neural network that has similarities to the mushroom bodies' structure.

3.2 The Auditory System

Positive phonotaxis in the cricket can be regarded as solving two related problems: recognising the calling song, in order to respond only to calling members of the same species, and locating the calling song, in order to move towards it. These functions are known to be achieved by a combination of auditory neurons in the nervous system and the physics of the ear system.

3.2.1 Directionality of the Peripheral Auditory System [Michelsen et al., 1994]

Auditory processing in the cricket begins with the physical reception of sound by the tympanal organs which are located about 12 mm apart, one on each leg of the prothoracic segment. It is at these sites that sound is converted into activity in the auditory nerves. The ability of a cricket to distinguish the direction of a calling song is enhanced by the way the tympani function as pressure difference receivers. The membranes move in response to both external pressure and pressure within a network of tracheal tubes as sketched in Figure 3.1. Each tympanal organ is connected by a tube (acoustic trachea) to an opening (acoustic spiracle) in the thorax on the same side of the body. The two acoustic tracheae are themselves connected across the mid-line of the body by a connecting trachea containing a central membrane. The ear system is therefore treated as having four inputs. Although there are actually two tympanal membranes
on each leg, laser vibrometry recordings show that under open-field sound stimulation the vibration amplitude of the posterior membrane exceeds that of the anterior by 10–25 dB [Larsen, 1987]. So, while the anterior membrane can be regarded as a sound input, its contribution is insignificant when the posterior membrane is functioning normally.

Figure 3.1: Schematic of the physical layout of the cricket ear. I – Ipsilateral to sound source, C – Contralateral to sound source. The tympani (T) are located on the foreleg tibiae. Each one is connected with the acoustic spiracle (S) on the same side of the body through a tube (acoustic trachea, ATI/ATC). The tracheae branch and join across the mid-line, forming the connecting trachea, ATX. This is divided by a central membrane (dotted line). Bold arrows show the three main contributions to the membrane response at Tl for 4.7 kHz sound [Michelsen et al., 1994]

When a cricket is standing at an angle other than head-on to a sound source, a small drop in sound volume will occur between the ipsilateral and contralateral sides of the body due to attenuation of the sound pressure as the wave spreads. There will also be a small effect due to diffraction around the cricket’s body. [Michelsen et al., 1994] records the maximum difference in sound amplitude between the two tympani to be at most 2 dB, when the sound source is between 60 and 90° to the animal’s mid-line. The amplitude difference between a tympanum and the opposite spiracle is similar. This amplitude difference was measured in close to ideal anechoic conditions. In natural environments with multiple sound sources and echoes this would be a poor cue for lateral localisation.

The phase of the sound wave reaching each ear also provides information about the sound source’s location. For a sound source directly in front of the cricket the
vibrations at the two tympani would be exactly in phase, but for a laterally placed sound source the phase of the vibration at the contralateral leg would be delayed relative to the ipsilateral side. Assuming a separation of 12 mm between the two tympani a sound wave of the frequency of the calling song (4.7 kHz) will take approximately 35 μs to travel between the two inputs when the cricket is side-on to the sound source. Diffraction again makes some contribution, so that the measured wave at the contralateral tympanum and spiracle may lag that at the ipsilateral tympanum by 90° ([Michelsen et al., 1994], Figure 4).

If the phase difference between the two sides is to be exploited by the cricket the two waves have to be compared. It is this comparison that is achieved by the tympani acting as pressure difference receivers. Sound reaching the side of the body contralateral to a sound source travels back through the connecting trachea to reach the inner surface of the ipsilateral tympanum. In the trachea the waveform is subjected to filtering which both attenuates and delays the signals in a frequency-dependent manner. At the frequency of the calling song the main cross-body contribution comes from the contralateral spiracle. Sound at the spiracle entrance is transmitted with a gain of 0.44 on average, whilst the contribution from the contralateral tympanum is negligible. Sound from the ipsilateral spiracle also makes a significant contribution. In fact for frequencies around 4.7 Hz sound coming via this route is amplified in the acoustic trachea.

A key contribution of Michelsen and colleagues' work [Michelsen et al., 1994] was to show how, when the vibrations in the three pathways (Figure 3.1) are summed at the tympanum the amplitude of the resulting vibration depends on the relative phase lags. The lag on each pathway consists of a fixed part due to transmission in the tracheae, and a variable part due to external arrival time and diffraction. As a cricket rotates with respect to a sound source these latter components of the lag change, creating a direction-dependent amplitude response. The contribution of the contralateral spiracle changes most rapidly in the frontal and posterior direction, resulting in a 5 dB difference between the ears when the sound source is at 30° to the mid-line. The phase of sound reaching the ipsilateral spiracle changes most around 90° and 270° to the mid-line and this results in a “contralateral null” where the response drops off rapidly as the sound source reaches 270°.

On the robot this enhancement of directionality is produced by an electronic circuit that uses delay lines to simulate the phase lags in the tracheae. The circuit is a cut-down version of a previous model [Torben-Nielsen et al., 2005], and uses the principle
of the pressure difference receiver with two inputs rather than four. It is described fully in section 6.4.1.

### 3.2.2 Auditory Neurons

#### Projection to the Prothoracic Ganglion

Vibrations at the tympani are transduced into electrical potentials in the primary afferent fibres of the auditory nerve. There are around 60 of these fibres per side, which travel along the leg and into the prothoracic ganglion. There they terminate close to the midline in an area known as the auditory neuropil [Imaizumi and Pollack, 2005]. In this region are also found arborisations of bilaterally paired neurons that show auditory responses [Wohlers and Huber, 1982]. Two identified neuron types with dendrites in this region are ascending neurons (AN1 and AN2) which send axons to the protocerebrum. Two further types are known as omega neurons (ON1 and ON2). These are local to the prothoracic ganglion, and have long curved processes that run between the auditory neuropil areas on the two sides of the midline.

#### Frequency Tuning of Prothoracic Neurons

Crickets use auditory cues during courtship behaviour and in the avoidance of echolocating predators, as well as in mate-finding. The sounds involved in the first two behaviours are of higher frequencies than the calling song, and the different types of auditory neurons reflect a division between the behavioural relevance of the different frequencies. AN1 can be classified as a low-frequency neuron, having a low threshold (\(\approx 40 \text{ dB} \)) for firing around the calling song frequency but a higher threshold of around \(70 \text{ dB} \) at \(2 \text{ kHz} \) or \(10 \text{ kHz} \). AN2 neurons tend to be most sensitive to higher frequencies (40 dB threshold at 14–16 kHz) but may also have a secondary region of less sensitive responsiveness around the calling song frequency [Schildberger, 1984b]. ON1 is tuned similarly to AN1 at low frequencies, with a secondary region of responsiveness centred on \(\approx 15 \text{ kHz} \) at 70 or 80 dB. ON2 neurons show broadband responses at moderate to high sound intensities [Wohlers and Huber, 1982].

#### Copying of the Chirp Structure

As described in Section 2.4.2 the most crucial parameter of the chirp pattern for determining the strength of a cricket’s response is the syllable repetition interval (SRI).
Neurons involved in the production of phonotactic turning ought therefore to show pattern selectivity that matches the behavioural response, or preserve the chirp pattern so that their post-synaptic targets may filter for relevant patterns.

None of the prothoracic neurons mentioned show a correlation between their firing rate and the behavioural response to different SRIs. AN1 and ON1 are, however, better at copying the chirp pattern than the other two types due to shorter membrane time-constants [Wohlers and Huber, 1982]. This suggests that they may relay the chirp pattern on to other brain regions that discriminate between more or less attractive songs. In contrast, longer time constants in AN2 and ON2 mean that syllables of the calling song tend to get blurred together. It should be noted that this effect is not entirely absent from AN1 / ON1, but shorter SRIs are necessary before it becomes noticeable. Since crickets do not respond phonotactically to these patterns it is possible that this blurring of high-frequency pulses forms part of the recognition process.

**Causal influence of Prothoracic Neurons on Phonotaxis**

On the basis of their frequency tuning and pattern preserving properties AN1 and ON1 are likely to be the most important of the four neurons for positive phonotaxis, although the involvement of AN2 and ON2 cannot be ruled out. This is supported by experiments in which an AN1, ON1 or AN2 neuron on one side of the body was hyperpolarised during phonotaxis to a standard calling song on an open-loop trackball [Schildberger and Hörner, 1988]. Hyperpolarisation of AN1 on the side most strongly excited by the calling song caused animals to turn to the opposite side, while the same intervention in ON1 caused a reduction in turning velocity. Hyperpolarising AN2 also caused a reduction in turning velocities in some animals but only at high stimulus intensities.

**Binaural Processing**

Excitatory input into the four pairs of prothoracic neurons mentioned above corresponds strongly with the distribution of their dendritic fields. In AN1 this is unilateral. In AN2 and ON1 the dendritic field is largely concentrated in the auditory neuropil on the side of the ascending axon, but a small region of arborisation on the opposite side also receives excitatory input. Only ON2 has binaural excitatory inputs of approximately equal size [Wohlers and Huber, 1982].

ON1 neurons are inhibitory. They are mutually monosynaptically connected, as
well as synapsing onto the AN2 neuron of the opposite side [Selverston et al., 1985]. Furthermore, cross-body inhibition of AN1 has been observed, with a tuning curve that is a strong match for the tuning curve for excitation of the opposite ON1, and similar latencies to the inhibition in ON1 and AN2 [Horseman and Huber, 1994]. It seems that the ON1 pair act to increase any inter-aural contrast in the neurons ascending to the brain.

**Filtering for SRI in Auditory Brain Neurons**

In the protocerebrum the ascending neurons arborise in regions which overlap with dendritic fields of certain local brain neurons. Synaptic connectivity has not been established directly, but the brain neurons show responses to sound with a short lag relative to the ascending neurons. It is therefore likely that they receive excitation from the prothoracic cells. These have been assigned to class BNC1. Further brain neurons showing auditory responses after a longer delay are assigned to class BNC2, which are assumed to receive input via the BNC1s, or even via other BNC2s [Schildberger, 1984b].

Anatomical separation of high and low frequency pathways, as seen in the ascending neurons, persists amongst the BNC1 and BNC2 neurons. Both classes contain examples with sensitivity profiles similar to AN1 or AN2, although the thresholds for firing are higher in every case.

Unlike the prothoracic auditory neurons the auditory brain neurons have been shown to be selective for different syllable repetition intervals. A member of the class BNC1 (BNC1d) was found to be a low-pass filter for SRI, and members of class BNC2 were found to act as high-pass filters or band-pass filters. The band pass filters showed a good match to the pattern selectivity shown by phonotactic crickets. This suggests that filtering for SRI takes place in the brain neurons, probably by at least one pair of neurons that combine individual high- and low-pass filtering characteristics to form a band-pass filter.

### 3.3 The Visual System

A great deal is known about visual processing in insects, although the majority of the information available comes from flies, bees and locusts. The small amount that is known from crickets is reviewed in [Honegger and Campan, 1989]. In more recent years the cricket has found favour only as a subject for investigations of polarised light
vision, which are largely irrelevant here. The gross structure of the cricket optic lobes has been revealed by staining [Honegger and Schürmann, 1975] and consists of the same series of layers as in other insects. From distal to proximal these are the retina, lamina, medulla and lobula. The expected function of each layer is outlined below, based on data from other insect species.

3.3.1 Structure and Function in the Optic Lobe

3.3.1.1 Retinal Resolution

The eye of the cricket is a compound eye with a number of facets that varies with species: around 3100 facets in *Gryllus campestris* [Honegger, 1980] and perhaps 7000 in *Gryllus bimaculatus* [Honegger and Campan, 1989]. An estimate of the interommatidial angle for *G. bimaculatus* can be obtained by assuming the eye is a hemisphere of unit radius and dividing the surface area by the number of ommatidia to obtain $2\pi/7000$. Assuming the curvature of an ommatidium is negligible the interommatidial angle is then approximately the angle at the apex of a right circular cone of unit height and base area $2\pi/7000$, that is $1.9^\circ$. It should be noted that an ommatidium does not necessarily sample a region of space subtended by exactly this angle. The acceptance angle may be somewhat larger, resulting in an overlap between adjacent ommatidia, effectively smoothing the image projected onto the retina.

3.3.1.2 Lamina

Photoreceptor cells from the retina synapse onto monopolar cells in the lamina. In flies the lamina monopolar cells' responses are inverted, amplified and more transient when compared to the photoreceptors [Laughlin, 1987]. This means there is no sustained response to a constant illumination level. Rather the cells act as high-pass filters, emphasising changes in intensity.

3.3.1.3 Medulla

The columnar organisation of the lamina continues into the medulla where repeated small circuits are expected to act as local motion detectors. The small size of the medulla cells makes recordings practically difficult, but those that have been successful have identified various motion-sensitive neurons (for a summary see the introduction to [Douglass and Strausfeld, 2003]). A recent model based upon a plau-
possible circuit of identified neurons has been shown to produce a qualitatively good fit to observed motion responses, while suggesting that the Hassenstein-Reichardt model (Section 2.8.2) is still a reasonable model for spatially-integrated motion responses [Higgins et al., 2004].

3.3.1.4 Lobula

The third neuropil of the insect optic lobe is known as the lobula. In flies, from which most information is known, there is an anatomical division of the lobula, the lobula plate, containing so-called tangential cells. These have large receptive fields, reflected in their wide dendritic arborisations over the retinotopic lobula complex. These cells integrate the output of local motion detectors over correspondingly wide areas of the visual field. Different types of tangential cell respond maximally to specific patterns of optical flow that correspond to particular manoeuvres in flight [Krapp et al., 1998]. Previously it had been suggested that a network involving the identified HSE, H1 and H2 cells might be organised to detect the type of wide-field yaw motion involved in the optomotor response [Krapp et al., 2001] although more recent work with naturalistic stimuli shows responses that contradict this prediction [Kern et al., 2005]. Nonetheless it is almost certain that the optomotor following observed in response to artificial stimuli results from wide-field integration in these types of cells, whether or not they are tuned to detect yaw specifically.

Whilst no tangential cells from the cricket have been studied, examples from the closely related locust have been characterised [Rind, 1990]. Two cells were identified with spiking responses to horizontal motion stimuli presented to the ipsilateral eye. On this basis it is reasonable to assume the presence of this type of cell in a model of the cricket.

3.4 Descending Neurons

Cricket brains contain around 300,000 neurons, but information passing from the brain to the ventral nerve cord is conducted in around two hundred bilateral pairs of descending neurons. This means that, for example, all information mediating visual control of walking and flight must be captured in the activity of this population.

A series of three studies in Gryllus bimaculatus has revealed some general principles about the functions of the descending neurons. Intracellular recordings were made
from the neck connectives of the crickets while they walked on an open-loop trackball. Stimuli of three modalities were presented:

- Auditory — *G. bimaculatus* calling song and 20 kHz chirps that elicited negative phonotaxis.

- Visual — moving gratings, and switching the lights on and off.

- Mechanical — air puffs to the cerci.

### 3.4.1 Separation of Modalities

An observation made in the earliest study of the three was that descending neurons that responded directionally tended to be unimodal [Böhm and Schildberger, 1992]. Whilst multimodal neurons that responded to all three types of stimuli were observed, they displayed rapidly habituating responses to all modalities, with no directional discrimination ([Böhm and Schildberger, 1992], Figure 5). Therefore those cells would be unable to control continuous steering behaviour. The authors suggest that they represent a kind of novelty detector. The two more recent studies report on a cluster of cells (named DBNi5) that show weak responses to optomotor stimuli whilst being characterised as primarily auditory [Staudacher and Schildberger, 1998] [Staudacher, 2001]. Apart from these, there are no cells reported with joint auditory and visual responses.

### 3.4.2 Maintenance of the Stimulus Structure

A feature of many of the neurons described is that their auditory responses copy the chirp structure of the cricket calling song (e.g. [Staudacher, 2001], Figures 6 & 7). This supports the idea, implied by the block diagrams of Chapter 2, of a direct transmission of sensory signals onto thoracic motor centres. Certain cells that showed a response to a moving grating displayed modulations of firing rate that were phase locked with the passage of stripes (e.g. [Staudacher and Schildberger, 1998] Figure 3B). It is possible that these cells were not motion sensitive but just responded to light intensity over a restricted receptive field. Other cells showed a clear directional response, and must have been responding to motion.
3.4.3 Possible Existence of a Population Code

It was noted by Staudacher that many of the neurons he recorded had partly overlapping stimulus/response characteristics. Many cells responded to the *G. bimaculatus* calling song, but some copied the chirp structure fully, some responded mainly to the onset of chirps, some were more responsive to ipsilateral stimulation and some were more responsive to contralateral stimulation. He suggested in [Staudacher, 2001] that this might indicate the use of a population code to increase the resolution and accuracy of information transfer.

3.4.4 Gating During Standing

[Staudacher and Schildberger, 1998] showed that the responses of descending neurons to calling song or optomotor stimuli were, in most cases, absent when a cricket stood still. This was not the case for air puffs to the cerci, ultrasound stimuli or switching the lights on and off. This seems to reflect the behavioural relevance of those stimuli, which can indicate the strike of a predator or the echo-location calls of a bat and presumably require a fast response under all circumstances. On the other hand positive phonotaxis and visual steering appear to require some inhibition to be released, depending on the behavioural context.

3.4.5 Correlation With Walking Parameters

All the studies found descending neurons that have firing rates that correlate significantly with some parameter of the cricket’s locomotion, e.g. forward walking [Böhm and Schildberger, 1992] (their Figure 6) or rotation [Staudacher and Schildberger, 1998] (their Figure 8). Staudacher reports finding such correlations for seven of nine neurons tested from the i5 cluster. He tested for, but did not discover, any neurons that were correlated with both translation and rotation. This suggests that the two parameters are encoded separately in the nervous system.

3.4.6 A Command Neuron that Initiates Walking

Command neurons are nerve cells, the spiking of which is necessary and sufficient to elicit some behaviour [Kupferman and Weiss, 1978]. One descending neuron in the cricket that perhaps fits this description was discovered by [Böhm and Schildberger, 1992]
(their Figure 7). Depolarising this cell was sufficient to elicit walking, and hyperpolarising it was sufficient to halt walking. This finding indicates that walking behaviour in crickets can be controlled from the brain.

3.4.7 Discussion

The descending neurons investigated by Böhm, Schildberger and Staudacher provide several pieces of evidence that are relevant to understanding and modelling co-ordination of multimodal behaviour. A key one for the current study is the idea that the visual and auditory modalities generally remained separated in the pathways leaving the brain. This supports the notion, implicit in all the block diagrams of Chapter 2, that the basic mode of combination is summation in the thoracic ganglia. It has to be assumed that any interaction between the pathways via corollary discharges must be physically located in the brain, occurring before sensory-to motor transformation. The studies reviewed did not generally use simultaneous presentation of different modalities, so such interactions would not have been revealed.

This separation of modalities is somewhat surprising, given the high numbers of multimodal neurons expected to be present in the central brain [Schildberger, 1984a] [Homberg, 2005]. It suggests that the multimodal neurons of the brain are likely to be dedicated to recognising contextual states, and using these to modify descending pathways [Wessnitzer and Webb, 2006]. The modality-specific behavioural gating described in Section 3.4.4 is presumably an instance of this.

3.5 Multisensory Processing in the Mushroom Bodies

Neural pathways from multiple sensory organs converge in the head ganglia of an insect. The fibre tracts are relatively easy to visualise, but the anatomy alone provides little conclusive evidence about how the signals they carry are combined, or what the consequences for behaviour might be. Electrophysiology, ablation studies, identification of mutants and various genetic intervention techniques have been used to clarify the functional roles of different neurons or structures.

Particular targets for these interventions are the mushroom bodies [Fahrbach, 2006]. These glomerular neuropil areas are known to have multimodal inputs and have a stereotypical organisation that presents the possibility of discovering general principles about multimodal integration. In contrast, non-glomerular regions of the brain
are likely to contain many multimodal neurons, but their structure is too disordered to make any reasonable progress towards understanding their function at a circuit level.

This section reviews some of the most important findings about the mushroom bodies and their implications for the co-ordination of the auditory and visual systems.

### 3.5.1 Morphology

The mushroom bodies (MBs) are a pair of symmetrical structures within the protocerebrum. Cells known as Kenyon cells are intrinsic to the mushroom bodies and give them their characteristic shape. Tightly-packed bundles of Kenyon cell axons form a stalk (pedunculus) running from dorso-caudal to rostro-ventral through the brain, which bifurcates to form two lobes, one pointing medially and one ventrally. Within the pedunculi and lobes bundles of Kenyon cells of different types can be distinguished by immunoreactivity, gene expression, cell body size or process morphology (reviewed in [Fahrbach, 2006]), but the functions of these heterogeneous cells types are not known. At the dorso-caudal end of each mushroom body the Kenyon cell dendrites form a pair of cup-shaped structures known as the calyces. Calyx morphology differs between species. In some insects there is only one per mushroom body, or they are almost non-existent. Bees and cockroaches have double calyces, but these are again morphologically distinct from the cricket’s.

Synapses between Kenyon cells and extrinsic neurons exist on all parts of the mushroom bodies, but most occur on the calyces and lobes. Based upon the form of processes seen in dye-filled neurons and the location of neurotransmitters and their receptors it appears that the calyces are input regions and the lobes are largely, but not uniquely, output regions [Fahrbach, 2006]. Efferent extrinsic neurons in the lobes appear to have dendritic fields outside as well as within the mushroom bodies and may interact with descending neurons [Li and Strausfeld, 1997].

Some extrinsic neurons run between the lobes and the calyces in the protocerebrocalycal tract, forming a feedback loop [Grunewald, 1999] [Schildberger, 1983]. Immunohistochemical techniques have revealed that GABA is likely to be the neurotransmitter used by these neurons in the bee [Bicker et al., 1985] and the cricket [Strambi et al., 1998]. Electron microscopy has demonstrated the existence of synapses between Kenyon cell axons within the pedunculus [Schürmann, 1987].
3.5.2 Behavioural Functions

Olfactory Learning

The most firmly established role of the mushroom bodies is in odour discrimination tasks. Wild-type *Drosophila* can learn to associate arbitrary odours with electric shock punishment whereas mushroom body mutants, or flies that have had their mushroom body formation halted by hydroxyurea treatment are impaired in their learning [Heisenberg et al., 1985] [de Belle and Heisenberg, 1994]. Similarly, proboscis extension in the bee can be conditioned by pairing of odours with a sugar reward [Okada et al., 2007]. This can be blocked by local cooling of the mushroom bodies or antennal lobes [Erber et al., 1980]. The most important indication that some component of memory is located in the mushroom bodies comes from a study in the *Drosophila* learning mutant *rutabaga*, which lacks the expression of an adenylyl cyclase involved in synaptic plasticity. Rescuing the expression of the adenylyl cyclase in the mushroom bodies only was sufficient to restore olfactory learning to wild-type levels [Zars et al., 2000].

Visual Learning

*Drosophila* with mushroom bodies ablated using the hydroxyurea technique have been tested in a variety of learning paradigms with non-olfactory conditioned stimuli. Both classical and operant conditioning using visual patterns and heat shock reinforcement were tested, as well as operant conditioning of yaw torque using heat shock [Wolf et al., 1998]. MB-ablated flies did not perform significantly worse than intact flies in any of these situations. However a more complex version of the visual operant conditioning task has been shown to require the mushroom bodies [Liu et al., 1999]. This particular variant of the task involves the introduction of a separate context variable (the colour of the background illumination, or the presence or absence of short “flashes” of darkness). Only flies with mushroom bodies could learn to avoid the reinforcement signal in one context and retain their pattern preference following a context switch.

Spatial Behaviour and Learning

A second type of experiment discussed in [Wolf et al., 1998] used heat shock to punish free-walking flies if they entered a particular half of a small arena. Both intact
and MB-ablated flies could be conditioned to avoid that part of the arena, whether in the presence of visual cues or in the dark. This is in contrast to a similar experiment in cockroaches, where the insects were placed in a heated arena containing an invisible cool spot. In the presence of visual cues on the arena walls intact cockroaches could learn to locate the cool spot progressively more quickly over several trials [Mizunami et al., 1998b]. Cockroaches with lesions to the pedunculus / β-lobes retained their initial performance level despite repeated training.

**Courtship Conditioning**

Effects of mushroom body ablation are also seen in courtship conditioning in *Drosophila*. Naïve intact male flies will attempt to court and mate with a virgin female fly, but will also unsuccessfully court a mated female or immature male. This unsuccessful courting results in a reduced response on subsequent presentation of a virgin female. MB-ablated flies, tested in the dark, do not undergo this conditioning. MB-ablated flies tested in the light initially display the same suppression of courting behaviour, but regain naïve levels after 30 minutes [McBride et al., 1999] (cited in [Zars, 2000]).

**Control of Locomotion**

Although flies with greater than ninety percent of their Kenyon cells absent have been described as “remarkably normal in many behaviours” [Heisenberg, 1998] a quantitative examination of walking in *Drosophila* has revealed that patterns of locomotion do change [Martin et al., 1998]. Ordinarily, flies walk in a series of bouts and pauses, with only short pauses occurring within bouts. Mushroom body ablated flies display longer bouts than intact flies.

### 3.5.3 Modality of Extrinsic Neurons

A major input to the calyces in most species including the cricket comes from the antennal lobe via the antennal-glomerular tract. This suggests that olfactory and mechanosensory input might dominate. In the honeybee there is also a major visual tract (anterior-superior optic tract) supplying the calyces [Mobbs, 1982] but this is not present in the cricket; staining of the cricket optic lobes [Honegger and Schürmann, 1975] revealed only a few fibres entering there. Other extrinsic neurons connect the calyces with many parts of the proto- and deutocerebrum,
suggesting that their inputs are likely to be multimodal. This was confirmed for the house cricket *Acheta domestica* by Schildberger [Schildberger, 1984a] who studied the responses of 322 extrinsic neurons to five types of stimuli: olfactory (banana odour), visual (constant white light directed at compound eyes), auditory tones and mechanical stimulation of the antennae or cerci. It was found that extrinsic neurons at both the lobes and calyces responded to all of these stimuli, with 58% showing responses to more than one modality. Responses to stimuli of many modalities have also been reported for the cockroach extrinsic neurons [Mizunami et al., 1998a], although the proportion of multimodal neurons was not made explicit.

For the purposes of the current investigation it is interesting to note the existence of a cell ([Schildberger, 1984a], Figure 2a) with an auditory response, apparently tuned to the calling song frequency. Although Schildberger’s investigation of auditory brain neurons [Schildberger, 1984b] did not reveal any that had terminals in the mushroom bodies this cell shows that auditory responses do occur there. Unfortunately the single sound pulses used do not reveal anything about the pattern selectivity of the cell.

Similarly, Schildberger’s use of simple flashes of light as visual stimuli mean there can be no certainty about whether visual motion information reaches the cricket mushroom bodies. Extrinsic neurons in the cockroach have been reported to respond to optical flow (for example [Li and Strausfeld, 1997], Figure 5 and [Mizunami et al., 1998a], Figure 6B, which responds regardless of direction), which may generalise to the cricket.

In the cricket the proportion of neurons responding to a given modality was similar in the calyces and lobes. Only olfactory responses differed significantly between the two regions, with 57% of calyx extrinsic neurons responding but only 20% of lobe neurons. It might be suggested that the non-olfactory modalities are present in the mushroom bodies only to provide context for olfactory learning. However if this were the case there seems no reason why lobe extrinsic neurons would respond to presentation of a non-olfactory stimulus alone.

### 3.5.4 Sustained Firing and Motor-Linked Activity

The most striking difference between the responses of calyx extrinsic neurons and lobe extrinsic neurons seen in the cricket [Schildberger, 1984a] was the presence of after-effects in the lobe. That is to say that a short stimulus could evoke elevated or depressed firing in a lobe extrinsic neuron that lasted for up to a minute. 28% of lobe extrinsic neurons showed these after effects compared to 8.3% in the calyces. Li and
Strausfeld notes a cell in the cockroach that responds to cercal stimulation with a burst of activity lasting around a second after the stimulus has ended ([Li and Strausfeld, 1997], Figure 9) and suggest that multiple input pathways with different delays or recurrent connections may be responsible for the effect. In principle either mechanism could explain the effect. However Schildberger notes that the average latency of a response to a sensory stimulus was 42.9 ms at the calyces and 68.4 ms at lobes. This lag of 25 ms is presumably due to a combination of axonal and synaptic delays in the Kenyon cells, and falls far short of the delay length needed to produce the observed effects.

In cockroaches the use of extracellular electrode arrays has allowed the activity of extrinsic neurons to be recorded in synchrony with film of the insects’ behaviour [Mizunami et al., 1998a] [Okada et al., 1999]. These experiments identified sensory neurons which appeared to reflect the activity of proprioceptors at the base of the antennae ([Mizunami et al., 1998a], Figure 11). An example of particular interest was a lobe neuron ([Mizunami et al., 1998a], Figure 10) that responded to external stimulation of the antennae but not to self-stimulation during grooming behaviour. This would appear to reflect the combination of direct sensory input with an efference copy.

The experiments also revealed lobe extrinsic neurons that displayed motor-related activity. In some cases the responses seen were attributable to proprioceptive sensing. Other cells ([Mizunami et al., 1998a], Figures 15–17) were classed as motor prediction units. These units would begin firing either spontaneously or in response to a sensory stimulus. Locomotion would be initiated after 1–2 seconds and the elevated firing rate would persist throughout the movement. In [Okada et al., 1999] it is suggested that these cells are copying the activity of descending neurons, i.e. they receive corollary discharges. Alternatively, the authors suggest, the extrinsic neurons might be directly involved in the formation of motor commands, but this is less likely when the lack of direct connections to thoracic motor centres is considered.

Due to the extracellular electrodes used in the recordings it was usual for multiple units to be recorded simultaneously. It was observed that physically proximal cells tended to display similar but not identical spiking behaviour [Okada et al., 1999]. This lead the authors to suggest that clusters of lobe extrinsic neurons might act jointly to transmit an MB output signal.
3.5.5 Mechanism of Learning

Mechanisms of learning in the mushroom bodies have, for the present, only been addressed in the context of the well established olfactory learning paradigms described in Section 3.5.2. These require that the signals mediating the conditioned stimulus (CS) and unconditioned stimulus / reinforcer (US) converge in the mushroom bodies. Correlated presentation of the two must result in subsequent presentation of the CS activating a neuron or population of neurons that leads to the conditioned response.

In both bees and fruit flies it has been shown that the reinforcing signals (US) are mediated by neurons that release either dopamine or octopamine (reviewed in [Giurfa, 2006]). Stimulation of an identified neuron in the bee brain, VUMmx1, can substitute for sugar reward in conditioning of the proboscis extension reflex. It projects to the calyces, antennal lobe and lateral horn, and injection of octopamine into the calyces or antennal lobe has the equivalent effect [Hammer and Menzel, 1998]. Imaging studies in Drosophila initially suggested that dopaminergic cells might mediate aversive reinforcement [Riemensperger et al., 2005]. Subsequent experiments in Drosophila larvae confirmed that octopamine and dopamine are sufficient substitutes for appetitive and aversive unconditioned stimuli, respectively [Schroll et al., 2006]. Pharmacological intervention has shown that the two amines have the same roles in crickets [Unoki et al., 2005].

While it has been demonstrated that plasticity within the Kenyon cells is necessary for olfactory conditioning (Section 3.5.2) a separate study has used reversible blocking of synaptic transmission in Kenyon cells to examine memory acquisition, storage and retrieval separately [Dubnau et al., 2001]. A temperature-sensitive transgene was expressed in the Kenyon cells, permitting synaptic output to be blocked by raising the temperature to 30°, and re-instated by reducing it again. It was found that synaptic output was only necessary during the retrieval stage of an olfactory association. A major implication of this finding is that the recurrent activity in the population of Kenyon cells and feedback neurons is not required for the acquisition or storage of an olfactory association.

The presence of learning despite a lack of Kenyon cell output leads Dubnau and colleagues to suggest that the synaptic changes underlying olfactory conditioning are likely to take place at the calyx rather than the lobes. This assumes that the mechanism of learning must be Hebbian learning based on the correlation of pre- and post-synaptic spikes. However, what is being learnt during olfactory association is in fact
a correlation between a sensory input and the reinforcement signal, which is likely to be mediated by a separate dopamine / octopamine signal. This means that learning at the Kenyon cell outputs could still take place if the plasticity was dependent only on a pre-synaptic correlation between Kenyon cell activity and the US.

An identified neuron (PE1) associated with the pedunculus in the bee has been shown to undergo a change in activity associated with conditioning of the proboscis extension reflex [Okada et al., 2007]. It shows reduced activity after conditioning. This lead the authors to suggest that PE1 exerts inhibition on descending, CR mediating, neurons before conditioning, which is subsequently released.

A finding superficially contradictory to those of [Dubnau et al., 2001] was reported in [Cassenaer and Laurent, 2007]: Hebbian spike timing-dependent plasticity (STDP) for the synapses from Kenyon cells onto β-lobe extrinsic neurons in the locust. However, the authors suggest that the role of this plasticity is not to learn weight changes, but to synchronise the timing of extrinsic neurons spikes with the oscillations of the local field potential. This is supported by the length of time over which the STDP has an effect. Co-incidence of spikes is thought to be an important feature of coding throughout the locust olfactory system [Perez-Orive et al., 2002].

### 3.5.6 Discussion

One of the propositions of this thesis is that the mushroom bodies of the insect brain could act as the predictive model component in an efference copy scheme like the one shown in Figure 2.13. This idea is put forward on the basis of the following reasoning:

- **Actions of the cricket that lead to re-afferent input can be initiated by sensorimotor pathways of many modalities.** Multiple sensory modalities including audition are represented in the presumed input areas of the mushroom bodies [Schildberger, 1984a].

- **Short discrete sensory inputs can initiate long sequences of behaviour and hence extended sequences of re-afferent input.** Presumed extrinsic output neurons of the mushroom body exhibit extended sequences of activity that far outlast an experimental stimulus [Schildberger, 1984a]. This is likely to be caused by recurrent activity.

- **In order to act as a processor of corollary discharges the mushroom bodies must receive input from one reflex-like pathway, and be able to influence descending
pathways. Some of the motor-linked activity observed by [Mizunami et al., 1998a] [Okada et al., 1999] could reflect either incoming corollary discharges or outgoing efference copy signals acting to modify the transmission of other sensory signals. That the mushroom bodies have an influence on walking activity was shown by [Martin et al., 1998].

- If the predictive model is adaptive, then the mushroom bodies must receive feedback error signals. It is known that biogenic amines may stand in for reinforcement in olfactory learning tasks [Giurfa, 2006]. Feedback error learning is closely related to operant conditioning. In operant conditioning the association of a motor command to a certain sensory and behavioural context becomes consolidated or abolished by positive or negative reinforcement. In feedback error learning of the type envisaged the association of a sensory prediction to a corollary discharge is modified by a feedback error signal which can be considered the analogue of negative reinforcement.

A number of the studies cited report experiments in which visual input replaces olfactory input as the conditioned stimulus in a learning paradigm. These have demonstrated that in most cases mushroom body ablation or mutation does not have the same detrimental effect on visual learning as in olfactory trials. The place memory experiment of [Mizunami et al., 1998b] is an exception. It should be noted that in the arrangement of Figure 2.13 the optomotor signal is not a direct input to the predictive component, but contributes to the feedback error signal. So visual input in the model does not have the same rôle as it has in the published experiments, and the absence of visual learning in most of those trials does not preclude the possibility that mushroom body output could modify visual responses as assumed.

### 3.6 Summary

This chapter has provided an overview of the cricket anatomy and neurophysiology relevant to the robot model of Chapter 6. The first two sections have set out the mechanisms known to be involved in the production of phonotaxis and optomotor following, and the model motor pathways of Chapter 6 are based closely on this experimental evidence.

Section 3.4 highlighted the main features of the signals transmitted in the neurons that descend from the brain to the thoracic and abdominal ganglia, and these are used
to justify the summation of auditory and visual modalities at the motor output, with interactions between the two taking place in sensory co-ordinates in the brain.

Section 3.5 reviewed the anatomy, physiology and behavioural roles of the mushroom bodies. These were used to justify the hypothesis that the mushroom bodies might act, amongst other things, as a centre for processing sensory predictions. The features that are thought to provide this capability are implemented at an abstract level in the predictive component of Section 6.3.
Chapter 4

Trackball Experiments

4.1 Introduction

As stated in Chapter 1, the aim of the work described in this dissertation was to try and identify how phonotaxis and optomotor following interact in the cricket. Five alternative hypotheses about possible mechanisms were defined in Section 2.6. The experiments of this chapter contribute both directly and indirectly to assessing the plausibility of the different schemes by testing crickets on an open-loop trackball.

Section 2.7.1 described an experiment which could, in principle, use open-loop stimuli to demonstrate the presence of a directionally-sensitive inhibition or gain modulation. The test was originally reported by [Böhm et al., 1991], but their results were inconclusive. The modelling work of [Webb and Reeve, 2003] shows that to infer the presence of such a mechanism requires demonstration of a significant reduction in the effect of an optomotor stimulus when a female cricket responds to a calling song from either side (Section 2.7.2). A replication of the original experiment is described in Section 4.4.1, which fails to show the presence of inhibition or gain modulation.

Having failed to find any evidence for these possible schemes, the chained subsystems hypothesis is addressed. As noted when the chained subsystems hypothesis was introduced in Section 2.6.3, the execution of auditory steering movements by the optomotor system requires that the optomotor system be capable of passing all the necessary signal frequencies. The experiments of [Hedwig and Poulet, 2004] have demonstrated rapidly alternating auditory steering movements. In Section 4.3.3 crickets are systematically tested with oscillating gratings to demonstrate that the optomotor system acts as a low-pass filter with a cut-off frequency of around 0.3 Hz. With the aid of a simple simulation described in Section 4.5.2 it is shown that the effect of this
filter on the auditory signal is not compatible with the amplitude of steering seen to split songs.

Neither the open-loop experiments of Section 4.4.1 nor the frequency response experiments of Section 4.3.3 are capable of distinguishing between the efference copy hypothesis and summation at the motor output (Sections 2.6.6 and 2.6.2). In Section 4.4.2 an experiment is described in which an optomotor stimulus is linked to the steering of the crickets, providing closed-loop visual feedback at the same time as open-loop auditory stimulation. This is used to present crickets with an alternating sequence of negative feedback, open-loop turning and positive feedback. The sign of the feedback had a significant effect on the magnitude of subsequent open-loop steering, indicating the presence of an adaptive mechanism. Section 4.5.4 considers where in the nervous system the adaptation is likely to take place, and explains why the effect seen is not compatible with the usual conception of an adaptive efference copy. An alternative mechanism is proposed to explain how a non-directional predictive model might produce the observed behaviour.

Certain trackball experiments contribute indirectly to the conclusions of the thesis by providing data to parameterise the robotic model. The optomotor system of the robot, described in Section 6.5.4, had its parameters tuned to match the behaviour of the cricket. This relied on the frequency response tests that have already been mentioned, measurements of responses to step changes in optical flow (Section 4.3.1), and measurements of the static characteristic curve of the optomotor response (Section 4.3.2). A small number of tests were also conducted with phonotaxis stimuli alone, in order to assess how turning on the trackball related to turning in the arena. These experiments are detailed in Appendix A.

### 4.2 Methods

#### 4.2.1 Insects

Female *Gryllus bimaculatus* from the colony maintained at the Department of Zoology at the University of Cambridge were isolated before their final moult and housed individually. They were fed on dried dog food and had water freely available. After emergence adults were located away from the sound of the colony in a laboratory where they were exposed to the local daylight variations and ambient temperature.

Before a cricket was used on the trackball a pin was attached to its back to allow
it to be handled and tethered in place. A small section of the fore-wings was removed and the pin was attached using a blob of wax on the anterior abdominal tergites. This arrangement was used in order to support the insect around its centre of mass.

4.2.2 Trackball

The trackball used was of the open-loop type introduced in Section 2.3. The layout of the components of the trackball and visual stimulus apparatus when in use are shown in Figure 4.1.

![Figure 4.1: Cross-sectional view of open-loop trackball and visual stimulation apparatus.](image)

The ball was 56.5 mm in diameter, made of a rigid plastic foam (Rohacell 31) and weighed 3 g. It fitted loosely into a transparent hemispherical cup, which was perforated with evenly-spaced holes and set into the top of a hollow plastic cylinder.
A constant air stream was passed into the cylinder, allowing the ball to float freely. Crickets were held above the ball by fixing the pin mentioned above into a needle holder attached to a 3-D manipulator. They were lowered so that their legs could turn the ball, and the air flow was adjusted so that the upward force approximated the weight of the insect under ordinary walking conditions.

Movements were registered by an optical mouse chip (ADNS-2051, Agilent Technologies) positioned within the air chamber underneath the ball. Motion was resolved into a forward–backward component and a left–right component (at right-angles to the body axis of the insects).

The trackball was fixed to the base of an anechoic box 150 cm wide, 100 cm deep and 70 cm high. Echoes from the walls, floor and ceiling were attenuated by black coloured acoustic foam (Illsonic Sonex 65/125, Illbruck, Germany).

4.2.3 Visual Stimuli

Wide-field optical flow stimuli in the horizontal direction were produced by surrounding crickets on the trackball with a patterned cylinder which could be rotated using a motor. The motor (Series 3564 Servomotor, Dr. Fritz Faulhaber GmbH.) was fixed into a metal plate suspended from the ceiling of the anechoic chamber on two rods.
4.2. Methods

These allowed the motor and cylinder to be raised for placing crickets on the ball, and lowered during experiments.

Patterned cylinders were made from paper backed with overhead projector acetate to provide some rigidity. When assembled they were 120 mm in diameter and 100 mm high. During experiments the bottom edge was at the eye level of the cricket which meant the moving pattern filled its field of view from the horizon up to an angular height of 60°.

Lighting was made as homogeneous as possible in the horizontal direction by using a circular fluorescent tube (22W, 50Hz AC) surrounding the patterned cylinder. This resulted in a lighter band around the centre as viewed by the cricket. Because a fluorescent tube cannot form a complete circle it was necessary for there to be a dark spot behind the animal.

According to Diagram 5C in [Honegger and Campan, 1989] the field of view for *Gryllus bimaculatus* extends dorso-caudally to around 120° and ventro-caudally to around 80° on the mid-sagittal line. The extent is even greater, up to 150°, elsewhere. This means that crickets walking within the cylinder would certainly have been able to see some of the acoustic foam lining of the anechoic box below its bottom edge. It is also possible that they could see some of the needle holder and manipulator arm, the trackball’s plinth, and the trackball itself. The ball had a speckled surface and would therefore produce the same optical flow cue that a cricket would see when walking on any similarly-patterned surface. The cricket’s body was fixed with respect to all other objects.

4.2.3.1 Control of Optical Flow Stimuli

The motor used to control the cylinder could be instructed to move continuously at a chosen speed, or a script could be used to schedule a sequence of movements and pauses. In addition the supplied controller could be used to simulate operation as a stepper motor. In the experiments of Section 4.4.2 this mode was used to create a feedback loop between the cricket’s walking and its visual input by using the left–right pulses of the trackball output to drive the pattern.

4.2.4 Auditory Stimuli

The auditory stimulus was always a synthesised *Gryllus bimaculatus* calling song with a carrier frequency of 4.7 kHz, 42 ms syllable repetition interval, 500 ms chirp interval,
6 syllables per chirp and 2 ms linear ramps at syllable starts and ends. The song was played back through at most one speaker at any instant. A number of speakers were located at fixed positions round an arc 87 cm from the trackball, allowing the angle of incidence of the sound to be switched. Most experiments used a pair of speakers (Sony SRS A57) positioned at 45° or 90° to the left and right, and the time-averaged multimodal tests (section 4.4.1) used an array of five small speakers at ±90°, ±45° and 0°. The volume of the sound was calibrated to 75 dB at the trackball.

It was unavoidable that the cylinder used to provide optical flow cues would interfere with the sound field around the crickets. Control experiments were run to test that the crickets’ ability to distinguish sound direction was still present after it was introduced (Section A.3).

### 4.2.5 Data Capture

The optical sensor of the trackball used quadrature encoding for each channel. This was converted electronically into a pulse-encoding with a 150 µs pulse being produced for each 127 µm of movement. Positive pulses signalled forward or leftward motion and negative pulses signalled backward or rightward motion, depending on the channel. Pulses were digitised at 10 kHz and stored to disk using a PC and data acquisition board (National Instruments PCI-MIO-16E-4) for later analysis.

Timings of auditory and visual stimuli were logged in synchrony with the trackball pulses using the same PC. The position of the patterned cylinder was registered by a potentiometer, connected to the drive shaft of the motor with a rubber belt. In addition, an optical sensor was used to detect the passing of the stripes of the pattern. Its output was 5 V when positioned over a black stripe and 0 V when over a white stripe. An RMS to DC converter (Analog Devices type 637) produced the envelope of the sound pattern.

### 4.2.6 Data Analysis

Initial data visualisation was done using the program Neurolab [Hedwig and Knepper, 1992], which is designed to perform common manipulations and graphing on biological time-series data. Instantaneous walking speed (referred to as “speed”) along both of the trackball’s component directions was calculated using the inverse of the recorded pulse intervals. The resulting traces were integrated to produce new series giving the cricket’s intended displacement along the
4.3 Optomotor Experiments

4.3.1 Step Response Tests

These experiments examined the optomotor response during step changes in optical flow. They were used in conjunction with the sinusoidal input tests of Section 4.3.3 to assess the lag and filtering characteristics of the optomotor system. This was important for correctly matching the dynamics of the robot visual system described in Section 6.5.4. A vertically striped cylinder with pattern period 120° was switched between zero and 42 r.p.m. (pattern frequency 2.1 Hz) as quickly as the acceleration of the motor would allow. Both directions of rotation were tested as detailed in Table 4.1. The particular rotation speed was chosen because it was the most effective speed in a pilot version of the characteristic curve tests (see Section 4.3.2). Four crickets were tested for two experiments of twenty minutes each, that is twenty repetitions of the protocol in Table 4.1.

<table>
<thead>
<tr>
<th>Time</th>
<th>0–15 s</th>
<th>15–30 s</th>
<th>30–45 s</th>
<th>45–60 s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pattern Frequency</td>
<td>2.1 Hz Left</td>
<td>0 Hz</td>
<td>2.1 Hz Right</td>
<td>0 Hz</td>
</tr>
</tbody>
</table>

Table 4.1: Protocol for step response tests

same two dimensions (referred to as “displacement”). Values could then be read off from the displacement traces at any chosen interval to obtain the average velocity along that direction. Sometimes it was useful to combine the orthogonal components and look at the length of the vector sum, which is referred to as “absolute speed” or “absolute displacement” as appropriate.

Further processing was done using Neurolab’s trigger-based averaging facility. Events of interest, for example chirps of the calling song, or stripes passing the optical sensor, were grouped together as trigger lists. These are lists of event times occurring on a certain data channel over possibly many files. Once a group of events were collected as a trigger list it was possible to examine the average value of any data channel or calculated trace during a chosen interval around the trigger times. An example is the average left–right speed over the duration of a chirp in Figure 4.16.

Statistical calculations were performed outside Neurolab, using either Matlab (Mathworks, Natick, MA.) or OpenOffice Calc (OpenOffice.org).
4.3.1.1 Results

Once placed on the trackball the crickets walked in a series of bouts and pauses. Average traces of the left–right speed component show that when the grating was still the crickets tended to walk straight, and when the motor was switched on they tended to turn in the direction of the grating in the frontal field (Figure 4.3, bottom).

![Graphs showing cricket movement](image)

**Figure 4.3:** Averaged forward–backward speed (top) and left–right speed (bottom) in response to step changes from 0 to 2.1 Hz in the left and right directions. Black lines show means, grey areas show standard deviations across four insects. Dashed lines separate periods of grating movement and non-movement. Horizontal line in bottom panels shows mean steering value over the full 60 s of the protocol.

During the first second following a step change in the grating speed the left–right acceleration of the crickets is very strong. Although there is certain to be some lag before the response it appears to be very short, and it is not possible to estimate it with any certainty given the level of noise. Subsequently the crickets’ average speed approaches an asymptotic value of around 1.5 cm s⁻¹. The time taken to attain half this maximum speed from the start of the stimulus can be read off from a zoomed version of the graph and is approximately 200 ms.
The shape of the curve may not be due entirely to the integration properties of the optomotor response because some crickets were observed to stand still in between movements of the grating, commencing walking some time after the motor started. An example of this situation is shown in Figure 4.4, during the second period of grating movement. The turning response of these insects will be delayed compared to that of those which are spontaneously walking when the stimulus begins. Because the average step response is comprised of a mixture of these two situations the slope will be determined to some extent by the distribution of times at which crickets start walking.

Figure 4.4: Response of a single cricket over one cycle of an optomotor step response experiment. Potentiometer (top trace) shows direction of cylinder rotation. Optical sensor (second trace) shows passage of stripes. Lower two traces show forward / backward and left / right displacement (note different scales). Short examples of non-walking during leftward grating movement are marked 'x'. Long period of non-walking overlapping the period of rightward grating movement marked 'NW'.

Forward–backward speed also changes throughout the step response (Figure 4.3, top). The average remains in the forward direction at all times. When the grating starts to move there is a drop in forward–backward speed, concurrent with the increase in left–right speed. When the two are combined to give the absolute speed it can be seen to remain roughly constant or slightly decreasing for about the first two and a half seconds of grating movement. This is consistent with the idea that rotation is independent of translation. Subsequently there is an increase in the forward component while the lateral component stays relatively steady. This could be interpreted either as
an increase in the speed of individual crickets, or the introduction of more crickets which are starting to walk.

4.3.1.1.1 Maintenance of optomotor turning over longer periods  It is noted in Appendix A that phonotaxis tends to reduce over time when a cricket is exposed to calling song for a long duration (see Figure A.5). Because the step-response tests involved presenting the same stimulus protocol repeatedly for twenty minutes they enabled a similar check to be made for the optomotor response. Figure 4.5 shows that the static phase of the step response showed no trend to decrease or increase in magnitude over the duration tested.

![](image)

Figure 4.5: Maintenance of step response magnitude over twenty minutes. Mean left–right speed during 15-second periods of grating movement. Averages of 4 animals ± standard deviation. Left: Grating movement in left direction. Right: Grating movement in right direction.

4.3.2 Characteristic Curve Tests

The aim of these experiments was to determine the magnitude of crickets’ turning tendencies in response to optical flow at different pattern velocities / temporal frequencies. This information was used to tune the time constant of the motion detectors used in the robot visual system (see Section 6.5.4). It is also necessary to have this data in order to correctly interpret the multimodal experiments of Section 4.4.1. Those experiments look for inhibition or gain modulation effects when auditory and optomotor stimuli are presented simultaneously. In order to recognise an increase in gain or a directionally-independent inhibition it is necessary to know the expected response to a unimodal optomotor stimulus.
Crickets walked on the trackball surrounded by a cylinder patterned with alternating black and white vertical stripes of period 120°. Ten-second periods of constant grating movement were interspersed with twenty-second periods when the grating was stationary. Pattern frequencies of 0.7, 2, 3, 4, 5 and 6 Hertz were presented in ascending and descending sweeps, as shown in Figure 4.6.

![Figure 4.6: Stimulus protocol for characteristic curve experiments](image)

### 4.3.2.1 Results

The magnitude of a cricket’s response to a particular optical flow stimulus was taken to be the average left–right speed over the ten-second period of stimulation. Figure 4.7 shows the characteristic curves obtained when the results are averaged across the six crickets tested, keeping the ascending and descending portions of the trials separate.

As expected, turning is in the direction of grating movement in the frontal field. Although no value was explicitly calculated for turning during the periods when the grating was stationary both curves pass close to (0,0) and are approximately symmetrical in other respects. As pattern frequency increases from zero the speed of turning increases at first then reaches a maximum before decreasing at higher frequencies. When the overall levels of turning from the ascending and descending sweeps are compared there is no significant difference ($t = 1.05$, $df = 5$, $p = 0.34$) so they are combined to produce the average curve shown in Figure 4.8.

Two refinements of the data are possible. Firstly, although the shape of the characteristic curve may be similar for all crickets certain ones have a generally higher level of locomotor activity and may skew the means with a relatively small sample. Secondly, lateral bias may be introduced by off-centre placement over the trackball,
Figure 4.7: Characteristic curves for the optomotor response at pattern frequencies up to 6 Hz, from ascending (solid line) and descending (dashed line) sweeps. Means of six crickets. Negative numbers indicate rightward grating movement in the frontal field (abscissa) and rightward walking (ordinate), consistent with the sign of the trackball pulses.

Figure 4.8: Characteristic curve for the optomotor response at pattern frequencies up to 6 Hz. Means of six crickets ± standard deviation. Ascending and descending sweeps combined.
whereas symmetrical behaviour is expected in response to symmetrical stimuli. In Figure 4.9 these factors have been adjusted for by subtraction of the difference between left and right stimuli, and normalisation by the overall activity level. This has had the effect of reducing the spread associated with each point and steepening the drop-off off the curve as the pattern frequency increases above 2 Hz.

Figure 4.9: Characteristic curve for the optomotor response at pattern frequencies up to 6 Hz (right). Means of bias-subtracted and normalised values for six crickets ± standard deviation. Ascending and descending sweeps combined.

4.3.3 Sinusoidal Input Tests

These tests aimed to characterise the frequency response of the optomotor system by testing the crickets with oscillating visual stimuli of different frequencies. These tests complemented the step response tests of Section 4.3.1; for a linear system it would theoretically be possible to obtain the same information from either type of test, but there was no reason to expect that the optomotor system should be linear.

As well as being important for the parameterisation of the robot model, this experiment was used in assessing the plausibility of the chained subsystem hypothesis (Section 2.6.3). In Section 4.5.2 the frequency characteristics discovered are built into a simple simulation which is used to reject the hypothesis on the basis that the optomotor filter would significantly attenuate the kinds of steering signals seen in split song tests (Section 2.7).

Sinusoidal oscillations with frequencies of 0.05, 0.1, 0.2, 0.4, 1, 2 and 4 Hz were used. The cylinder had a black-and-white vertically striped pattern as used in previous experiments, but with a smaller pattern period of 60°. It was hoped that the use of
narrower stripes would reduce the tendency of the crickets to fixate a particular part of the pattern, although there is some evidence that this failed (see last paragraph of results). The positional range of motion was fixed at forty-five degrees in all cases. Seven animals were tested.

4.3.3.1 Results

Figure 4.10 shows the stimulus function and response for frequencies of 1 and 2 Hz. The upper traces show the grating’s speed. It should be noted that the decision to fix the angular extent of pattern motion required that the grating move at higher instantaneous speeds at higher oscillation frequencies. This had to be taken into account in the analysis of response amplitude (see below). The lower traces show the mean left–right speed across all seven crickets, averaged using trigger points placed at every two cycles of the stimulus function. The response traces are dominated by the frequency of the stimulus, meaning it is possible to characterise them by their amplitude and phase lag with respect to the input. To quantify these properties the Fourier transform of the two-cycle average was calculated for each oscillation frequency. The component at the stimulus frequency is shown superimposed in the lower traces of figure 4.10. Similarly reasonable fits were obtained by this method in all cases.

Figure 4.10: Mean response across all crickets (N=7) to grating oscillating at 1 Hz and 2 Hz. Upper traces — grating speed. Lower traces — Mean speed of crickets (grey lines) with the dominant component of Fourier transform superimposed (black lines).

On examining Figure 4.10 it is noticeable that the crickets’ steering movements had
a smaller amplitude in response to the 2 Hz oscillation, despite the stimulus containing higher pattern frequencies than in the 1 Hz case. In order to interpret this observation it is necessary to consider the characteristic curve found in Section 4.3.2. This curve is non-linear and non-monotonic, so the responses to sinusoidal inputs with increasing ranges of pattern speed will not change in a straightforward manner. This is illustrated in Figure 4.11 for inputs of 0.4 and 1 Hz. To produce the figure Matlab’s fminsearch function was used to fit the experimental characteristic curve with a log-normal function: 

\[ f(x; \mu, \sigma) = \frac{K}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}} \]

with \( K = 9, \mu = 1.541 \) and \( \sigma = 1.026 \). This was applied to the input functions which are plotted with time on the ordinate. The outputs show the patterns of turning speed that would be expected assuming that the characteristic curve were produced by a zeroth-order process — that is a functional transformation with no time-dependence and hence no filtering.

For lower absolute pattern frequencies up to 1 Hz the characteristic curve might be reasonably approximated by a straight line. For inputs within this range a sinusoidal input is transformed into another sinusoid (Figure 4.11, left) and inputs of increasing amplitude produce outputs of increasing amplitude. When the range of input speeds overlaps the nonlinear parts of the characteristic the output changes; sinusoidal shapes become distorted (Figure 4.11, right), and the peak steering speed of the output is limited to the peak value of the characteristic.

In order to assess whether the observed amplitude changes between different stimulus conditions could be attributed to this process alone equivalent output curves were calculated for all input frequencies. Response amplitude was quantified as the mean absolute steering speed. This measure was used rather than Fourier analysis because it was known in advance that the outputs at high frequencies would not be true sine waves. In Figure 4.12 (left) these theoretical values have been plotted alongside the curve obtained from the crickets’ responses.

While the curves match for the lowest frequencies tested it can be seen that the amplitude of the crickets’ responses are much reduced from the expected level at higher frequencies. Only the very first point of the graph, corresponding to 0.05 Hz does not fit the trend. As there is no reason to believe that the system could cause an amplitude gain this can be regarded as an outlier. In the right hand graph of Figure 4.12 the data is shown as a Bode magnitude plot without this first point. Although it is noisy, this plot is comparable to that which would be expected for a first-order low pass filter with a cut-off frequency between 0.2 and 0.4 Hz [Cruse, 2006].

Examining the phase lags of the crickets’ responses provides complementary ev-
Figure 4.11: Visualising the effect of the nonlinear characteristic curve. Left — 0.4 Hz input. Right — 1 Hz input. Lower frequency oscillations require a low range of pattern speeds and sinusoidal input may be transmitted with little distortion. Higher ranges of input speeds result in noticeable harmonic distortion.

Figure 4.12: Left: Theoretical amplitude response at oscillation frequencies from 0.05 to 4 Hz assuming no filtering effect (dashed line) and measured response from crickets (solid line). Right: Attenuation relative to theoretical values, in dB. Response at 0.05 Hz omitted.
idence about the nature of the filtering taking place. This quantity can be measured as the delay between peaks of the upper and lower curves in Figure 4.10 (1 and 2 Hz only), or found directly from the Fourier transform. When plotted against stimulus frequency as shown in Figure 4.13 (left) the points seem to form a smooth curve. It is not, however, the curve that would be expected from a first-order low-pass filter, which would have an asymptotic phase lag of \( \pi/2 \) rad. The theoretical lag for a first-order low-pass filter with cut-off frequency 0.3 Hz is superimposed (dashed line). The discrepancy can be explained if a short fixed delay is introduced into the system as shown in Figure 4.13 (right).

![Figure 4.13: Phase lag, in radians, of crickets’ turning relative to grating for oscillation frequencies from 0.05 to 4 Hz (solid lines). Left: Overlaid with \(- \arctan(f/0.3)\) (dashed line). Right: Overlaid with \(- \arctan(f/0.3)\) plus a 76 ms fixed delay (dashed line).](image)

Two other features of Figure 4.10 should be mentioned. Firstly there are clearly higher-frequency oscillations of about 10 Hz superimposed on the stimulus frequency. Although this is faster than has been previously reported for the stepping cycle of the animal, it is difficult to suggest what else it might represent. The second observation of possible interest is that the turning responses, for high frequencies in particular, are biased to the left-hand side. This effect could be explained if the crickets had a tendency to try and turn towards a particular part of the grating pattern. For example there may have been a black stripe in their left frontal field of view (the grating moved less than one pattern period during oscillation).
4.4 Multimodal Experiments

4.4.1 Time-Averaged Responses

This experiment replicated the one described in [Böhm et al., 1991], and discussed in detail in Sections 2.7.1 and 2.7.2. The position of a speaker broadcasting calling song was moved around tethered crickets to produce a characteristic curve for phonotaxis. This was repeated in three different visual conditions: with leftward optical flow, rightward optical flow, or a stationary grating. The visual pattern used was the black and white striped grating of period 60°, which was rotated at 10 r.p.m. to give a pattern frequency of 1 Hz when required. Sound stimulation was done using the synthetic calling song described in Section 4.2.4. The protocol for a single trial is shown in Table 4.2. Nine crickets were tested for two or three trials each.

<table>
<thead>
<tr>
<th>Block 1</th>
<th>Grating</th>
<th>1 Hz right</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sound</td>
<td>90L</td>
<td>90R</td>
</tr>
<tr>
<td></td>
<td>45L</td>
<td>45R</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Block 2</th>
<th>Grating</th>
<th>Stationary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sound</td>
<td>90L</td>
<td>90R</td>
</tr>
<tr>
<td></td>
<td>45L</td>
<td>45R</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Block 3</th>
<th>Grating</th>
<th>1 Hz left</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sound</td>
<td>90L</td>
<td>90R</td>
</tr>
<tr>
<td></td>
<td>45L</td>
<td>45R</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 4.2: Stimulus protocol for time-averaged tests. A complete trial consists of three blocks, each with a continuous visual stimulus. The auditory stimulus changes direction as shown e.g. speaker at 45 degrees to the right is shown as 45R. '-' denotes a ten-second pause. All periods of sound stimulation are 30 s long.

As discussed in Chapter 2, Section 2.7.2, the test could have two kinds of outcome. Either the three curves representing the three visual conditions could be parallel, or the sound direction could influence the magnitude of the optomotor effect. The latter would be consistent with a directionally-dependent inhibition of the optomotor system, or a modulation of the optomotor gain, whilst the former would be compatible with any of the remaining three hypotheses (Section 2.6.7), or a non-directionally-specific inhibition or gain modulation. This last case can be checked for with reference to the characteristic curve calculated in Section 4.3.2.
4.4.1.1 Results

Average left-right turning values were calculated for each cricket in each stimulus situation, excluding the ten-second silent periods. These were then averaged across all animals to produce the graph in Figure 4.14 (left).

Figure 4.14: Left: Left-right walking speeds for nine crickets stimulated with calling song and grating combined. Means ± standard error. Right: Distributions of shift between lower and upper curves for each sound direction.

The central line of the left-hand graph is a (sparsely sampled) characteristic curve for open-loop turning to sound within the stationary visual environment. The other two lines, obtained when the grating is switched on, retain approximately the same shape but are shifted away from the central one according to the direction of the optomotor stimulus.

It is noticeable that the turning of the insects was biased to the left in all situations. It is most likely that this is due to a consistent error in the tethering of the insects over the trackball\(^1\). It cannot be attributed to the left speakers being louder than the right, as the shift remains even when the active speaker is to the insects’ front. Neither is it attributable to the insects being attracted to a particular part of the striped grating, since the bias is still present when the grating is moving. For these reasons the central point of the graph is treated as though it represents a neutral turning tendency, and the other values are examined relative to it.

The quantity of interest in figure 4.14 (left) is the size of the gap between the central curve and the “shifted” curves produced when grating movement is present. If the

\(^1\)Equivalently, it is also possible that the optical chip had become mis-aligned relative to the manipulator arm.
inhibition hypothesis or the gain modulation hypothesis were to hold (see Section 2.6), then the size of the shift might alter for different speaker angles. For example, if the cricket inhibits its optomotor response only when turning to sound, then the size of the shift should be reduced for lateral positions of the speaker compared to the frontal position. The same effect could be achieved by a modulation of the optomotor gain. The chained sub-systems hypothesis, efference copy and summation at the motor output are all expected to produce a parallel shift [Webb and Reeve, 2003].

In order to test the hypothesis that the size of the shift altered depending on the direction of the sound a one-way repeated measures ANOVA was conducted. The central curve of the Figure 4.14 (left) was unexpectedly asymmetrical, so rather than trying to examine the two shifts separately (lower to middle curve & middle to upper curve) the combined shift from the lower to upper curve was examined. The distribution of this shift for each speaker direction is shown in Figure 4.14 (right). Table 4.3 gives the ANOVA summary table.

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>Prob&gt;F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between conditions</td>
<td>2.119</td>
<td>4</td>
<td>0.530</td>
<td>2.1</td>
<td>0.1038</td>
</tr>
<tr>
<td>Between subjects</td>
<td>17.100</td>
<td>8</td>
<td>2.138</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>8.073</td>
<td>32</td>
<td>0.252</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>27.293</td>
<td>44</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.3: ANOVA table for one way repeated measures ANOVA comparing optomotor-induced shift for five speaker directions.

No significant effect of speaker direction was found, so the null hypothesis (that the speaker direction makes no difference to the influence of the optomotor stimulus) cannot be rejected.

An alternative possibility is that there is inhibition, but it is non-directional. In this case the optomotor response would be inhibited for all speaker angles. This can be tested by comparing the observed shift against the effect of an optomotor stimulus in the absence of sound. Having established that the shifts at the different speaker angles are not significantly different they can be combined to obtain an average shift of 0.91 cm/s (SD across crickets = 0.33). This is not significantly smaller than the shift (0.98 cm/s) that would be predicted by interpolation from the characteristic curve of Figure 4.8 (unpaired t-test (2-tailed), t = -0.5981, df = 8, p = 0.57). So there is no
evidence that the presence of sound reduced the effect of the optomotor stimulus.

### 4.4.2 Feedback Experiments

These experiments examined the effect of closing the visual feedback loop, by coupling the movement of the visual surround to the lateral movements of the trackball. In the basic configuration the feedback direction was set to simulate the visual flow that a cricket would experience when walking in an arena. For example, if the cricket were to attempt to turn right the trackball would rotate anti-clockwise as viewed from above, and the grating would be rotated left in the frontal field (see Figure 4.15, left). This is referred to as negative feedback. The opposite situation, in which the sign of the grating movement is reversed, is referred to as positive feedback (Figure 4.15, right). The crickets experienced periods of visual feedback, interspersed with equal periods when the grating was stationary, that is to say the feedback loop was opened again.

![Negative visual feedback](image1.png) ![Positive visual feedback](image2.png)

**Figure 4.15:** Overhead schematic of trackball showing movement of cricket, trackball and grating under positive and negative feedback conditions.

In order to stimulate the crickets to turn the calling song was played from $90^\circ$ to the left of the animal as described for the phonotaxis-only experiments (Appendix A). It was expected that the majority of turns initiated by the animal would be towards this sound source.

The aim of the experiment was to look for adaptation, which might indicate the presence of a predictive model. When a cricket is placed on the trackball then, assuming it has an adaptive predictive model, that model will reflect the properties of the cricket’s environmental action up until that time. In theory, as the feedback condition is altered throughout the experiment the cricket’s brain will detect the discrepancy between its prediction and its new feedback, and adjust the model to fit. According to the block diagram of Figure 2.13, a changed prediction should have a direct effect on
the insect’s walking. By using three different feedback conditions (positive, negative, open-loop) it is possible to examine the effect of two of the conditions on walking in the third (in this case open-loop). If the behaviour in the third condition shows a dependency on the previous feedback state, then there is a good cause to claim that feedback-dependent adaptation is taking place.

The grating with a pattern period of 60° was used and the feedback strength was set so that approximately 27 cm of lateral deviation resulted in one revolution of the cylinder. This meant that for a left-right steering speed of 0.7 cm/s, which was typical for phonotaxis in the stationary striped drum (Section 4.4.1), the rotation speed would be 9.33 °/s and the pattern frequency would be 0.15 Hz. Although this is relatively slow compared to the pattern speeds tested in other experiments it meant that the typical rotation speeds would be within the range experienced in the arena experiments (see Appendix C).

The exact protocol used for the tests is shown in Table 4.4. The eighty-second protocol was repeated between twelve and thirty times per trial (16 - 40 minutes). Twelve insects underwent at least two trials each, and one extra insect was tested once. It should be noted that the intention was for the periods of feedback to alternate sign, but a problem with the control of the motor meant that this did not work properly in most trials. The sign of the feedback tended to alternate, but occasionally one or two changes of sign did not occur. The selection of positive or negative feedback could not be said to be random.

<table>
<thead>
<tr>
<th>Time</th>
<th>15 s</th>
<th>5 s</th>
<th>15 s</th>
<th>5 s</th>
<th>15 s</th>
<th>5 s</th>
<th>15 s</th>
<th>5 s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sound</td>
<td>On</td>
<td>-</td>
<td>On</td>
<td>-</td>
<td>On</td>
<td>-</td>
<td>On</td>
<td>-</td>
</tr>
<tr>
<td>Feedback</td>
<td>On</td>
<td>Off</td>
<td>On</td>
<td>Off</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.4: Protocol for feedback tests. Total cycle of 80 s is divided into four blocks which alternate between “feedback on” (either negative or positive feedback) and “feedback off” (grating stationary).

4.4.2.1 Results

The recorded data were initially examined at the time-scale of a chirp (500 ms). Trials were divided up into blocks of positive feedback, negative feedback or open loop. Open-loop blocks were further designated as “post-positive” or “post-negative”, de-
pending on the sign of the preceding feedback. Neurolab's triggering facility was used to mark the start of each chirp during each of the four conditions. It was then possible to obtain the average left-right speed over a 500 ms chirp period for each insect in each condition.

When the averages across all insects are calculated the curves in Figure 4.16 (left) are obtained. The shape of the curve is similar for all conditions. A distinct steering movement towards the speaker commences around 60 ms after the start of a chirp, reaches a peak at around 300 ms then declines again until the response to the next chirp begins.

Figure 4.16: (Left: Mean L–R speed over one chirp period during positive, negative, post-positive and post-negative feedback blocks. +: Positive, < + >: Post-positive, -: Negative, < – >: Post-negative. Curves are means of data from 13 animals. Right: Integrals of curves for all individuals in all conditions.

The obvious difference between the curves is that they are separated out in a particular order. The curve corresponding to negative feedback displays the weakest turning towards the speaker and the curve for positive feedback shows the strongest turning. The separation of the middle two curves is surprising because they were obtained from identical stimulus conditions. If these curves are truly different then it must be due to the preceding feedback state.

For statistical comparison the integrals under the four curves were calculated for all crickets separately. The distributions of scores in the four feedback conditions are shown on the right of the figure. It can now be seen that the ordering of the four curves is a trend displayed across many crickets. The difference between the post-positive and post-negative conditions is confirmed by the Wilcoxon signed ranks test ($T = 8, p$
4.5 Discussion

4.5.1 Characterising the Optomotor System

The gross structure of the cricket optic lobe is the same as is seen in flies and other insects [Honegger and Schürrmann, 1975] [Egelhaaf, 2005], suggesting that the same mechanisms of motion detection could be present. However, even if this is the case, it is reasonable to expect that differences in behaviour between crickets and flies will have resulted in the evolution of a somewhat different tuning of the temporal properties of the system.

The static characteristic curve of Section 4.3.2 supports the first assertion by showing a log-normal curve shape typical of the Hassenstein-Reichardt model of motion detection [Srinivasan et al., 1999b]. The second assertion is borne out by the fact that the cricket’s visual system is maximally sensitive to a pattern frequency of around 2 Hz; this is half of the value seen for Drosophila, on which the optomotor system of the previous robot crickets was based.

For the purposes of robot building it is required that the new visual system is a better match to the target animal. By using the standard Hassenstein-Reichardt model of elementary motion detection the required static characteristic can be arrived at by tuning of the low-pass filters that create the delays. Unfortunately the multiplication in the model introduces non-linearity which makes it hard to say how the particular filtering characteristics measured in the sinusoidal input tests and step response tests should be arrived at.

The sinusoidal input tests produced output that looks reasonably like a first-order low pass filter with a characteristic frequency of 0.3 Hz (equivalently $\tau =530$ ms). A first attempt at a model based on this data might neglect the temporal characteristics of the EMDs, and treat them as a zeroth order process with the given characteristic curve feeding into such a low-pass filter. This is not a good match for the step response tests, however. There, assuming a first order linear filter, the measured half-time of 200 ms would suggest a shorter time constant, of $\tau =290$ ms. Therefore it is not possible to neglect the filtering contribution of the EMD model. Instead it is necessary to be aware of the temporal characteristics of the EMDs, and possibly add further filters to get a reasonable fit to the data. This was done by simulation on the robot, as described in
4.5.2 Chained Sub-systems Hypothesis

This hypothesis (section 2.6.3) suggests that phonotaxis may be carried out in the presence of an optomotor response if the output of the phonotaxis system feeds into the optomotor pathway. Turns to sound would then be carried out as if they were turns to optical flow. Consequently, injected phonotaxis signals would be subject to the filtering properties of the optomotor pathway.

From an anatomical perspective it is highly unlikely that the phonotaxis system influences the optomotor system at the level of transduction, in the way that the haltere system of flies is affected physically by visual input [Chan et al., 1998]. Assuming that the overall optomotor response is produced by the contributions of multiple layers in a feed-forward network, the exact nature of the filtering applied to phonotaxis signals would depend on where in this network they were injected. At one extreme summation at the motor output (Figure 2.9) might be considered as an example of chaining: the phonotaxis signals are injected at the point where the optomotor pathway synapses onto neural networks that co-ordinate the walking pattern. The filtering applied to them is then due only to these walking networks, and the inertia of the cricket's body. This somewhat absurd example emphasises that the summation and chained subsystem models are closely related, and that the phonotaxis signals have to enter the optomotor pathway early enough in the chained sub-systems scheme for it to be distinct.

It has already been established that the complete optomotor pathway acts as a low-pass filter with a cut-off frequency around 0.3 Hz (Section 4.3.3). For current purposes it is assumed that the phonotaxis signals would enter the optomotor pathway early enough that the filtering they undergo is not significantly different from this. Can anything be concluded about the plausibility of the chained subsystems hypothesis based on whether the auditory system also has this property?

In the strictest sense it is not possible to accept or reject the chained sub-systems hypothesis on this basis, because it is not known how sound signals might be transformed in the brain before the auditory and optomotor pathways converge. Auditory signals might feed into the optomotor pathway, but be filtered in a different way overall due to a currently undiscovered intermediate filter. Conversely, if phonotaxis does show the same filtering properties then it could be due to an independent low-pass filter acting in parallel to the optomotor system. If both filters were linear filters then this
would be functionally equivalent to the chained subsystems scheme, but the assumptions of anatomical convergence would not be met.

It is only possible to infer anything about the plausibility of the chaining hypothesis if it is assumed that the auditory signals are not modified beyond what is observed in the ascending AN1 neurons or local brain neurons (Section 3.2.2). That is to say that pulses of sound heard by the insect are transmitted to the brain as bursts of spikes matching the pulse duration, or as single spikes. In this case it is possible to ask what output we would expect to see for different auditory inputs, given what we know about the optomotor pathway.

The usual procedure for investigating the frequency response of a system involves presenting a range of different sinusoidal inputs, as was done for optomotor system. However, this cannot be done for auditory steering because crickets only turn to patterns close to that of the conspecific calling song. Split-song paradigms, for example see [Hedwig and Poulet, 2005], have demonstrated that crickets will turn to individual syllables of calling song played alternately from two speakers on opposite sides of the body. This retains the structure of the song while inducing rapidly alternating steering at approximately 12 Hz. This is probably the upper limit on the frequency of phonotactic steering that could be induced in an experiment. A unilateral calling song stimulus as used in this chapter induces modulations in turning with a frequency of 2 Hz.

Due to the square-shaped pulses present in the calling song pattern the waveform contains a range of frequencies and is not simply amplified or attenuated by a linear low-pass filter. Figure 4.17 shows the effect of filtering the standard calling song pattern or a split song pattern with first-order low-pass filters of characteristic frequencies 0.3 Hz, 1 Hz and 3 Hz.

Filtering the standard song produces an output with a DC-bias that increases for lower cut-off frequencies. Pulse- and chirp-correlated modulations decrease as the cut-off frequency is reduced. The same kind of variation in pulse-correlated modulation is seen in the filtered split songs.

Inspection of the averaged turning tendency of crickets in response to a split song ([Hedwig and Poulet, 2005], Figure 8D) reveals a roughly sinusoidal modulation of turning tendency with a peak-peak amplitude of around 1.3 cm/s. Figure 7A from the same paper shows average turning to unilateral calling song, with a peak-trough amplitude of 1.7 cm/s. The ratio of these amplitudes (1.7/1.3 ≈ 1.3) gives an idea of the goodness of fit of the different filters in figure 4.17. At the experimentally-determined cut-off frequency of the optomotor filter the chirp-correlated modulations
Figure 4.17: Effect of filtering on standard and split calling song pattern. Input patterns consist of chirps with six square wave pulses of 21 ms each and 21 ms pulse intervals. Chirp repetition interval is 500 ms. Filters with cutoff frequencies of 0.3 Hz, 1 Hz and 3 Hz shown.
have an amplitude of 0.126 (no units) and the pulse modulations of the split-song response have amplitude 0.041. The resulting ratio ($\approx 3.1$) is much higher than for the cricket. As the cut-off frequency is increased the ratio decreases to 2.5 at 1 Hz and 1.3 at 3 Hz.

On the basis of the simulation, assuming that there is no intermediate filter on the auditory pathway, it is reasonable to say that the chained subsystems hypothesis does not hold, but that phonotactic turning may be subject to low-pass filtering with a higher cut-off frequency than that of the optomotor system.

### 4.5.3 Rejection of the Inhibition Hypothesis

As stated at the start of the chapter, the original experiment of [Böhm et al., 1991] might have permitted an inhibitory mode of combination to be identified, but too few crickets were tested to be sure. The time-averaged tests of this chapter replicated the original experiment with nine insects instead of three, and failed to reveal the predicted reduction in the effect of a moving grating for lateral speaker positions. The shift between the characteristic curves for the "grating still" condition and the "grating right" condition did appear to be reduced for speaker positions to the left of the cricket, but this was paired with a significantly increased shift for leftward grating movement. Moreover, the corresponding effect was not seen for symmetrical stimulus conditions with the speaker on the animals' right. This lead to the conclusion that the observed variation in shift was due to asymmetry in the curve for the "grating still" condition.

If this explanation is accepted then the shift between the upper and lower curves of Figure 4.14 (left) is the best indicator of the magnitude of the optical flow-induced effect. As the box plots in the right of the figure show, there is a slight tendency for grating movement to induce a larger shift in turning for lateral positions of the speaker. This would contradict the predictions of the inhibition hypothesis.

It should be noted that the magnitude of any reduction in shift would depend on the time course of inhibition within the nervous system. For example, the auditory signal that ascends to the brain in the spike patterns of the AN1 neurons copies the calling song pattern closely [Wohlers and Huber, 1982]. If it is assumed that the inhibitory signal acts to completely block the optomotor system with this exact time course then the inhibition effect would only be present for 288 ms in every second, that is to say it would reduce the observed shift by only 29%. Given the variability among the animals it might not be possible to conclude anything about an effect of this size.
The feedback tests offer some assurance that this is not the case by demonstrating a significant response to optical flow even when grating movement is restricted to those times when the cricket is turning and hence, by definition, when the inhibitory signal would be acting.

4.5.4 Feedback Experiments

The surprising result of the feedback experiments is the discovery that the open loop behaviour of the crickets changes depending on the preceding feedback situation. This suggests that some component of the system adapts during stimulation with visual feedback.

None of the five mechanisms outlined in Chapter 2 necessarily contain any adaptive components by definition, but any of the components involved in the different mechanisms could be adaptive. The phonotaxis component is the least likely source of the change. In the previously published literature there is no discussion of plasticity in phonotaxis besides a variation in the preferred sound pattern linked to temperature change [Doherty, 1985]. In Appendix A some evidence for a decrease in turning speed over a the first few minutes of a trial is presented in Section A.4. However this does not suggest anything other than a time-dependent decrease in motivation. The idea of short term plasticity in visual systems and predictive systems is better supported.

4.5.4.1 Adaptation in the Visual System

Adaptation is observable in the fly’s lobula plate tangential cells in response to stimulation with sustained optical flow [Harris et al., 2000]. The static phase of the step response to a test stimulus of moderate velocity is reduced after presentation of a high velocity adapting stimulus. If such adaptation can occur in crickets it was not expected that it would have affected the responses during the open loop phase of the feedback tests, since the grating was stationary. However, the mechanism of tethering did mean it was possible for the crickets to move their heads from side to side. It is possible to envisage how, during feedback, the optomotor following response might have adapted in such a way that it became asymmetric during the subsequent open-loop phase of the experiment. Oscillatory head movements by the cricket might then induce a bias in turning that would appear superimposed on turns to sound.

At present there is no evidence to suggest whether or not such adaptation might take place in the cricket. The repeated testing of step responses reported in Section 4.3.1
does not show any adaptation, but the pattern frequency being tested there was ten times lower than that used in the fly experiments of [Harris et al., 2000]. If high pattern frequencies are necessary to induce adaptation then the grating speeds reached in the feedback tests are unlikely to have been sufficient.

Could a simpler explanation suffice? As demonstrated by the step response tests (Figure 4.3), the left-right steering of the cricket does not change instantaneously, but builds up or falls away over the course of a few hundred milliseconds when a moving grating is started or stopped. In the feedback tests the crickets received self-generated optomotor input which stopped when the test entered an open-loop phase. Would the residual activity in the optomotor system be enough to account for the gap seen between the closed-loop curves?

The difference in turning speed between the negative and positive feedback conditions in Figure 4.16 was 0.84 cm/s. Under the assumption that the auditory and optomotor systems are summed at the motor output this must be due to the difference in optomotor input, since the auditory stimuli are identical. The difference between the two open-loop curves is 0.22 cm/s. In the step response tests the crickets attained an average left-right speed of 1.35 cm/s during the grating movement (2.1 Hz). During the fifteen-second periods without grating movement that followed the average left-right speed was 0.15 cm/s, or 11% of the speed during stimulation. On this basis it seems unlikely that the same kind of residual activity could account for the difference between the two open-loop conditions in the feedback tests.

4.5.4.2 Adaptation in a Forward Model

Efference-copy type mechanisms are commonly assumed to involve an adaptive forward model, as has been shown to be the case in electric fish [Bell, 2001]. However the results of the feedback tests are not entirely consistent with the usual conception of an adaptive efference copy. The block diagram of Figure 4.18 shows how a general adaptive efference copy mechanism would be expected to work in the cricket under negative feedback conditions.

The angular position of the speaker relative to the cricket, \( \theta \), is detected by the ears. Taking anticlockwise angles to be positive, a speaker positioned on the left of the cricket causes a positive signal, \( \alpha \) to pass through the phonotaxis system, resulting in anticlockwise turning. Throughout the discussion that follows it is assumed that \( \theta \) stays constant and positive, as in the experiment.

The lower part of the diagram shows the optomotor controller, which involves a
negative feedback loop. Anticlockwise turning (positive $\phi$) is transmitted as a negative value to the optomotor system ($\omega < 0$). In the absence of the predictive model this signal would act in the opposite direction to $\alpha$.

Assuming that the predictive model component has learnt the correct function to calculate the reafference the efference copy $\alpha$ is transformed into a prediction $\pi$ so that $\omega + \pi$ is always zero in the absence of external optical flow ($\phi_{ext} = 0$). Under negative feedback conditions a stable prediction $\pi$ should have the same sign as $\alpha$.

If the feedback loop is broken, as it is in the open-loop phases of the experiment, $\omega$ becomes zero, and $\omega + \pi \neq 0$. The output of the predictive model is summed with $\alpha$, altering the cricket's turning tendency. At the same time the error signal (dashed line) causes the predictive model to adapt its output towards zero. It is under these circumstances that the cricket's behaviour under open loop conditions could vary, according to the form of the prediction.

If this model is consistent with the results of the feedback experiments on the trackball it should be the case that the predictive signal is positive after a period of positive feedback, and negative after a period of negative feedback. This is not the case. Under negative feedback $\omega$ is negative and $\pi$ is driven towards positive values. Under positive feedback $\omega$ is positive and $\pi$ is driven towards negative values. In this case the opposite ordering of the two open-loop curves would be observed. The results of the feedback
tests are not compatible with the formulation of efference copy presented here.

### 4.5.4.3 Proposed Alternative Form of Predictive Signal

It is still possible for the results of the feedback tests to be explained by some kind of efference copy mechanism, if the form of the predictive signal is different. The standard efference copy mechanism as described above assumes that the predictive signal necessarily captures both the magnitude and direction of the expected re-afference. In nature, there is no reason why a cricket should ever experience positive feedback from its optomotor system. Therefore, there is no reason why an efference copy mechanism must capture the sign of the feedback.

Instead of the scheme previously described, suppose that the predictive signal is always positive, capturing only the magnitude of the expected re-afference. The absolute value of the incoming optomotor signal is subtracted from this prediction, resulting in a positive signal if the animal needs to turn faster, or negative if it needs to turn slower. This would serve the purpose of compensating for small disturbances in turning, assuming there is never a reversal of direction.

In a negative feedback situation the prediction would adapt to a stable value, whereas in a positive feedback situation the size of the prediction would grow large to try and compensate for the high levels of optical flow caused by self-sustaining feedback. This would lead to a higher level of expected optical flow after the positive feedback situation, resulting in enhanced turning.

### 4.6 Summary

Replication of the time-averaged multimodal tests of [Böhm et al., 1991] failed to produce the pattern that would be expected from a directionally-sensitive inhibition mechanism (see Section 2.7.2). Comparison of the effect of the optomotor stimulus during sound stimulation to the effect of optomotor stimulation alone (Section 4.3.2) also found no evidence for outright inhibition or gain enhancement.

A parallel curve shift in the experiment of Section 4.4.1 could be a product of any of the other mechanisms listed in Section 2.6.7. Establishing the frequency response of the optomotor system provided good evidence to suggest that the chained sub-systems hypothesis should be rejected, but the experiments of this chapter do not permit a choice to be made between the two remaining hypotheses (summation at the motor
output and efference copy). The feedback tests revealed an unexpected dependence of open loop turning behaviour on a preceding feedback situation. This shows that some form of adaptation is taking place in the system, possibly in a forward model. The results of the feedback tests are not, however, compatible with the usual form of the efference copy model. A modified version of an efference-copy based mechanism which takes into account only the magnitude of the response was proposed, and was tested by robotic modelling, as described in Chapters 6 and 7.

The trackball experiments also provided new information to permit a temporally-scaled model of the cricket’s optomotor behaviour to be produced. This differs in its parameters from the model based on *Drosophila* that was used on previous cricket robots, and permits a more meaningful comparison between the robot and the cricket.
5.1 Introduction

Experiments with the trackball system of Chapter 4 showed that summation at the motor output or efference copy are the most likely of the five hypotheses introduced in Section 2.6 to explain the crickets’ behaviour. It was not possible to make a decisive choice between them, however. The tests with reversed feedback indicate the presence of feedback-dependent adaptation, but this doesn’t have the characteristics expected of a directional efference copy (Section 4.5.4).

The tests described in this chapter investigate further the possibility of an adaptive model, by testing crickets in fully closed-loop conditions in an arena. It was hypothesised that if the crickets were not inhibiting their optomotor responses during phonotaxis (Section 4.5.3), and did not use an adaptive predictive model, then altering their visual environment should have an observable effect on their walking. This can be seen by reference to the block diagrams of Figures 2.9 and 2.13; increasing or decreasing the contrast of the visual surroundings has a direct effect on the stage labelled “environmental filtering”. On the other hand, the absence of an effect would indicate that the crickets were managing to compensate for the change. This would support the idea of an adaptive prediction.

Assessment of the behaviour of crickets during free walking has been used extensively in phonotaxis research in the past (see Section 2.3.2), as well as in assays of visual responses [Lambin, 1984]. A range of approaches to behavioural analysis can be used, from simply scoring whether the cricket reaches the speaker or not [Pollack and Hoy, 1979], to detailed assessments including the position, body angle and even leg position [Murphey and Zaretsky, 1972]. In this study path descriptions
based on position only were used to compare different visual conditions. Appendix C describes some additional experiments in which the arena was used to film phonotaxis in more detail, using a hand-held camera.

Crickets were tested in three visual conditions: with blank arena walls for low visual contrast; with high contrast stripes added; under infra-red light which the crickets cannot detect, and hence should have the same effect as dark conditions. As described below, it was not possible to detect any significant differences between the crickets’ walking paths in any of the pairs of conditions. This is more consistent with the idea that they adapt to anticipate the motion feedback from different visual surroundings.

### 5.2 Methods

#### 5.2.1 Experimental Arena

The arena was enclosed in a wooden box constructed for the purpose of attenuating external sound and internal echoes. It was intended that the arena would be used for testing the robot as well as crickets, so that both would experience the same environment. The arena floor was made as large as possible whilst still allowing the whole space to be viewed by a wide-angle camera mounted on the box’s ceiling. The box was 122 cm high, 175 cm long and 122 cm wide, with the two smaller side panels hinged for access. The floor was made of 25 mm chipboard, and the sides, walls and roof were made of 18 mm medium-density fibreboard. The arena could be lit by a circular fluorescent tube (22 W, 50 Hz A.C.) mounted in the centre of the ceiling. During later experiments some pieces of translucent, white-coloured plastic were mounted underneath the light to act as a diffuser.

The inside of the walls and roof were covered in white acoustic foam tiles (Illasonic Audio 90/125 Illtec, Illbruck, Germany). The base was uncovered and formed the floor of the arena, where the crickets would walk. A thin coating of clear varnish was applied to the floor to permit the surface to be wiped clean, and this was roughened slightly.

During experiments crickets were constrained to walk within a certain rectangular portion of the floor by 23 cm high walls placed within the box. The walls had a melamine resin coating so the crickets could not climb them, and were white in colour. With the internal walls in place the arena floor measured 97 cm by 128 cm. For trials involving sound one of the shorter internal walls was replaced with a piece of white,
acoustically transparent cloth. A pair of speakers (Samsung SMS-2200) were placed behind it, one at each end. Crickets walking on the arena floor could hear the speakers but could not locate them visually. The layout of the inside of the arena as used in all the cricket experiments is shown in Figure 5.1. An overhead view of the arena taken by the webcam in the box ceiling can be seen in Figure 5.2.

![Diagram of the arena layout](image)

**Figure 5.1:** The layout of the arena. Speakers were placed as shown for cricket experiments, but were moved for robot experiments. Details in Chapter 7

### 5.2.2 Sound Stimuli

All experiments used the same synthesised *Gryllus bimaculatus* calling song described for the trackball experiments of Chapter 4: carrier frequency of 4.7 kHz, 42 ms syllable repetition interval, 500 ms chirp interval, 6 syllables per chirp and 2 ms linear ramps at syllable starts and ends. Whilst the arena contained two speakers, only one speaker was active during any given trial. The volume of the sound was calibrated to between 70 and 75 dB using a sound-level meter (Dawe D-1422C / CEL 254) placed at the centre of the arena and angled towards the active speaker.

### 5.2.3 Visual Stimuli

The visual surroundings of crickets walking in the arena were altered by the introduction of dark-coloured stripes or landmarks to the arena walls. Specific arrangements are detailed along with the relevant experiments.
5.2.4 Insects

Female *Gryllus bimaculatus* from commercially bred stocks were isolated before their final moult and housed individually, away from the sound of male crickets. They were fed on dried dog food, and had water freely available. Heat and light were provided by ordinary tungsten light bulbs on a 12 hr:12 hr cycle, although the light cycle was lengthened by the local natural daylight in summer. Day time temperatures varied between 22 °C and 28 °C. Crickets were used for experiments between 7 and 43 days after the final moult unless otherwise stated.

5.2.5 Experimental Procedure

Experiments were always conducted as repeated measures tests, that is the same insects were tested in multiple conditions. For a given batch of insects it was most usual for only two conditions to be compared. To reduce the influence of order effects counterbalancing was employed; the whole set of insects would be divided into two age-matched groups, which differed in which experimental condition they experienced first. Individual crickets were tested as follows:

1. Cricket taken from its box on a piece of egg box, which was placed in the centre of the arena.

2. Arena closed and video recording started

3. Sound started after five seconds of video recording

4. Sound and video recording terminated when cricket reaches the edge of the arena OR after four minutes of inactivity

Crickets which remained inactive throughout their first test in a given condition were retained and re-tested at a later time. In addition to the possibility of a trial being rejected if the cricket did not respond, it would also be terminated if the cricket started exploring the arena before the sound was started.

Throughout an experiment the choice of which speaker was broadcasting was varied from cricket to cricket. When possible this was done on every other trial so that crickets would not be able to associate the presence or absence of an olfactory cue with the path to the speaker.
5.2.6 Data Capture

To accurately record the positions of crickets throughout an entire trial a fixed camera (Webcam Live Ultra, Creative Labs) was mounted in the centre of the box's ceiling, pointing through the middle of the fluorescent tube. Its wide-angle lens viewed the whole arena floor. Recordings were made at a resolution of 640 x 480 pixels and a frame rate of 15 frames per second. In images captured by the webcam crickets appeared as dark dots against the lighter coloured floor. No information about the insects' body orientations could reliably be extracted from these images.

5.2.7 Tracking

Individual video frames were extracted from the saved videos using MPEG decoding software (TMPGEnc, TMPG Inc., Gardena, CA., U.S.A.). Tracking of the crickets throughout the resulting sequences of frames was performed using semi-automated tracking software [Rosano, 2007]. The software was first calibrated to correct for barrel distortion caused by the wide-angle lens. The initial location of a cricket was marked manually, then the software automatically tracked its subsequent positions using a probabilistic algorithm based on colour matching. It was occasionally necessary to resume marking the cricket's position manually if it walked into a shadowed area, walked next to a black object or was obscured for some reason.

5.2.8 Path Properties

Once a cricket had been tracked throughout a sequence the following properties of its path were extracted using MATLAB scripts.

- Direct path length, indirect path length, and their ratio.
- Track time, mean speed and instantaneous speed profile.
- Instantaneous headings, normalised mean vector and D-statistic.

Path Length

In a typical phonotaxis experiment a cricket would leave the piece of egg box on which it had been placed, and walk in a meandering path towards the active speaker. On reaching the edge of the arena it might find the cloth wall and climb it. In many cases
the cricket would stay on the floor and walk around the perimeter of the arena (wall-following behaviour). For the purposes of analysis the end of such a path was defined as the last position of the cricket that was greater than 2.5 cm from the edge of the arena floor. This ensured that the analysed paths did not contain segments of wall-following behaviour.

The start of a path was defined to be the first position of the cricket lying outside a given circular area at the centre of the arena. Because the crickets tended to start off hiding under their piece of egg box the exact position at which they first became visible varied. For a given experimental condition the starting circle was made just large enough that all the analysed paths started at the same distance from the arena centre.

Once the start and end point of a path were determined a simple indication of the spatial directness was calculated. The direct path length is the straight line distance between the start and end points. The indirect path length is the sum of all the straight line segments composing the tracked path. The ratio Direct/Indirect takes values from one, for a completely straight run to the speaker, towards zero for increasingly meandering paths.

**Normalised Mean Vector and the D - Statistic**

A mean vector [Batschelet, 1981] is used to express the mean and variance of a group of angular quantities, taking into account that angular scales are cyclic. The mean angle is given by the vector direction. The variance is given by the length, which decreases from one, when all angles are equal, to zero when the angles are evenly distributed around the circle.

In the analysis of these behavioural experiments the angles in question are those of each straight line path segment with respect to the direction of the sound source. The normalised mean vector used here [Webb and Reeve, 2003] additionally takes into account the fact that not all path segments are of equal size. The contribution of the angle of each segment is weighted according to its length, while retaining the normalisation of the mean vector's length. The angle and magnitude are then interpreted as for the standard mean vector.

The normalised mean vector is calculated as follows: The position of the cricket in video frame $i$ is denoted $(x_i, y_i)$, and it is assumed that the speaker is at $(0, 0)$. A vector
is calculated for each successive pair of positions with

\[ \text{distance}_i = \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2} \]

and

\[ \text{heading}_i = \arctan \left( \frac{y_i}{x_i} \right) - \arctan \left( \frac{y_i - y_{i+1}}{x_i - x_{i+1}} \right) \]

The sum \( \text{length} = \sum \text{distance}_i \) gives the indirect path length. The magnitude and angle of the normalised mean vector are then calculated using

\[
\bar{x} = \frac{\sum \text{distance}_i \cdot \cos(\text{heading}_i)}{\text{length}} \quad \bar{y} = -\frac{\sum \text{distance}_i \cdot \sin(\text{heading}_i)}{\text{length}}
\]

\[ \text{magnitude} = \sqrt{\bar{x}^2 + \bar{y}^2} \quad \text{angle} = \arctan \left( \frac{\bar{y}}{\bar{x}} \right) \]

Once the normalised mean vector has been obtained it can be used in the calculation of the directness statistic \( D \):

\[ D = \text{magnitude} \cdot \cos(\text{angle}) \cdot \text{tracktime} \]

This is a product of three quantities which cause it to tend to one in the case of fast, straight, directed paths, and zero in the case of slow, mis-directed, meandering paths. \textit{Tracktime} is constrained to the interval \([0, 1]\) by definition:

\[ \text{tracktime} = \frac{\text{min.time}}{\text{actual.time}} \]

\textit{min.time} is taken to be the time it would take the cricket to run the direct path length at 15 cm \( \cdot \) s\(^{-1} \), which is faster than the average speed maintained by any of the crickets tested.

### 5.2.9 Classifying Phonotaxis

The intention was to assess the effect of the visual surroundings on phonotactic walking. Over all, given that the crickets generally approached the active speaker, it is reasonable to say that the calling song had an effect, and that phonotaxis was being displayed. However, there is no way of saying for sure whether a given cricket was being influenced by the sound on a single trial. In order to avoid assessing the path properties of many trials which were not influenced by the sound a rough criterion was applied: only those crickets that reached the arena wall in the quadrant containing the active speaker were assessed. It is still possible that some crickets may have arrived in
the correct quadrant purely by chance, but these cannot be separated out from the true sound trackers. For some visual situations the occurrence of non-phonotactic walking is of interest, so in many cases the results are presented twice, both with and without the non-trackers.

5.3 Phonotaxis With and Without a Grating

In these experiments the paths of crickets were compared in the presence and absence of a black-and-white grating pattern on the walls of the arena. The striped patterns were made of cardboard, and were made the right size that a cricket standing at the centre of the arena would experience the same spatial frequencies along the horizon as a cricket on the trackball at the centre of the six-striped drum (see Section 4.2.3). After these conditions had been compared a third one was tested, in which the black stripes were replaced with graded stripes that faded from black in the centre to white at the edges. Figure 5.2 shows overhead photographs of the arena, demonstrating the different types of stripe.

![Figure 5.2: Left – Arena with solid stripes. Right – Arena with graded stripes](image)

5.3.1 Taxis Without Grating

Eleven crickets were tested. At the start of the experiments the youngest cricket was eleven days post-moult, and the oldest was thirty-five days post-moult. The crickets were split into two groups (Group A – ages 35, 24, 20, 20, 14 and 12 days post-moult. Group B – ages 35, 20, 20, 13 and 11 days post-moult). Group A experienced the striped condition first while group B experienced the blank condition first.
5.3. Phonotaxis With and Without a Grating

Walking paths were recorded for ten of the eleven crickets tested in the blank arena. The remaining cricket was tested four times and always remained inactive. Of the responsive crickets, nine walked on their first trial, and one required two re-tests. Figure 5.3 (left) shows the paths of the ten crickets which walked.

![Figure 5.3](image.png)

**Figure 5.3:** Paths in the arena with the grating absent. Speaker positions are marked by the small squares at the right-hand end. Left – First run for all ten crickets which walked. Right – As left but with one path which finishes in an incorrect quadrant removed.

Nine of the ten tracks show a typical meandering course towards whichever speaker was active. One of them contains a small loop, but the cricket still approached the correct speaker afterwards. Only one track is discounted from consideration because it ends in the wrong quadrant of the arena, closer to the speaker in the upper part of the diagram rather than the lower one, which was active at the time. In Figure 5.3 (right) that track has been removed.

5.3.2 Taxis With Black Stripes

When the black stripes were present all crickets walked on every trial, whether or not the behaviour was classifiable as phonotaxis. The response of each cricket to its first trial in the presence of the grating is plotted in Figure 5.4. From visual inspection it appears that the stripes caused a change in the crickets’ behaviour; whilst all except one still headed towards the end of the arena that concealed the speakers the spread of the finishing positions was greater. Three of the eleven crickets reached the wall in a quadrant of the arena other than that containing the active speaker and so are not considered to have displayed phonotaxis. These tracks have been removed in the right hand side of the figure.
The tracks that remain still appear to be more spread out than those of figure 5.3. In particular some tracks reach the wall near the edges of the black stripes rather than near the speaker. On this basis it was hypothesised that the crickets were being influenced by the presence of the stripes, but not in the way originally expected. Rather than using the high-contrast edges to stabilise their phonotactic course it seemed that the crickets might be responding to the stripes as distinct landmarks, by walking towards them.

5.3.3 Taxis With Graded Stripes

In order to try and separate out any stabilising effect of increasing contrast from the object responses thought to be produced by the black stripes the experiment was run again, using graded stripes as pictured in Figure 5.2 (right). The death of some of the crickets from the previous experiment and their replacement with new ones meant both parts of the experiment were repeated to allow the use of repeated measures statistics. Group 1 contained 6 crickets (28, 28, 22, 20, 15 and 11 days post-moult) which experienced the striped condition first. Group 2 contained 6 crickets (43, 28, 28, 21, 19 and 15 days post-moult) which experienced the blank condition first.

In the blank arena paths were recorded for eight of the twelve crickets, all of which approached the active speaker on their first test (Figure 5.5, left). The remaining four were re-tested once each, but remained inactive. In the arena with graded stripes ten of the twelve crickets walked, nine on their first test and one after three tests. One track finished in an incorrect quadrant of the arena, and has been removed from the
5.3. Phonotaxis With and Without a Grating

plot (Figure 5.5, right).

Figure 5.5: Paths in the arena with blank walls (left) and graded stripes (right). The black rectangles show the positions of the cardboard pieces that carry the graded stripes.

It is difficult, based on the plots, to say whether the use of graded rather than solid black stripes made any difference to the crickets’ tendency to walk towards them. Certainly in the lower part of the right-hand diagram there are still two tracks that approach the centre (the darkest part) of one of the stripes.

5.3.4 Comparing the Three Conditions

Ideally it would be possible to say whether changing the visual environment by introducing stripes makes any difference to the directness of phonotaxis. However, the apparent tendency of crickets to approach the dark stripes means it is difficult to say whether paths recorded in the striped conditions show phonotaxis, a kind of scototaxis behaviour, or a combination of both.

Across all trials it is reasonable to say that the calling song has an influence. If it did not then the path end-points would be distributed more evenly around the arena perimeter. It might be argued that the crickets have a preference for corners, or that they can detect some visual difference between the fabric wall and the solid walls, and prefer to go to the fabric one. Even so one would expect the tracks to arrive indiscriminately at one or the other end of the fabric wall. In fact, for the blank arena tests nine out of ten crickets arrived in the correct quadrant on their first test. For the solid stripe tests the proportion is eight out of eleven, and for the graded stripe tests it is nine out of ten.

Only six of the crickets tested in the first experiment walked in both the blank and
the solid stripe conditions. If their D scores are plotted as in Figure 5.6 (left) it is apparent that the scores tend to be slightly higher in the striped case (0.36 compared to 0.23), and also that the scores become more spread out ($\sigma^2 = 0.026$ compared to $\sigma^2 = 0.0026$). The difference in medians is not significant (Wilcoxon matched pairs test, $T = 2, n = 6, p = 0.094, \text{ns}$), however the increase in variance is ($F$ ratio test, $F = 10.01, df = 5, 5, p = 0.024$).

![Figure 5.6](image)

**Figure 5.6**: Left — D scores for crickets that walked in both the solid stripe and blank conditions. Right — D scores for crickets that walked in both the graded stripe and blank conditions.

When the three components of the D statistic are examined separately it turns out the majority of the change is actually due to the *tracktime* component which assesses how fast the crickets walked. The cosines of the mean vector directions reduce slightly on average when the stripes are present; this relates directly to the spread of the path endpoints away from the speakers. The mean vector length increases slightly, which corresponds to less variance in the heading of the insects, regardless of what point they are attempting to approach. It should be remembered, however, than none of these trends are significant.

In the graded stripe tests eight crickets walked in both conditions. Their D scores are plotted in Figure 5.6 (right). Once again there is no significant change in the median (0.34 in the blank condition to 0.30 with the graded stripes). The variance is larger in the striped condition ($\sigma^2 = 0.022$ compared to $\sigma^2 = 0.014$) but this is not significant ($F$ ratio test, $F = 1.52, df = 7, 7, p = 0.59$).
5.4 Phonotaxis In Visible Light vs. Infra-Red Light

An alternative method of comparing phonotactic paths under conditions of varying optical flow was to test the crickets in the light and the dark. The "light" condition was exactly the same as the blank condition from Section 5.3.1. Even in the absence of black or graded stripes the crickets would receive optical flow cues from the ridges in the acoustic foam that lined the arena. Other possible cues to movement would have come from the corners of the arena, where the shape of the brown floor meeting the white walls would have caused some horizontal optical flow.

In the dark no optical flow can be perceived. In crickets there is not expected to be any sensitivity to light at the red end of the spectrum, so an infra-red light source was used to permit the tracks to be recorded by the overhead CCD camera without affecting the insects' sensory input.

As for the previous experiments crickets were separated into two groups of six crickets, which differed in the order in which they experienced the conditions. The ages of the crickets were seven to eighteen days post-moult in the first group and eight to nineteen days post-moult in the second. The experiment was conducted as a rapid pilot test and it was not considered necessary to alternate the side of the speaker between trials (see tracks in Figure 5.7).

Overall the crickets walked with a lower probability in this experiment compared to the tests of Sections 5.3.1 to 5.3.3. Due to the low levels of walking the crickets were each tested multiple times in order to try and get enough data to allow a statistical comparison between the two conditions. At the end of the experiment each cricket had been tested four times in the light and three times in the dark. Out of the sixty trials thirty-one were terminated because the cricket remained inactive after four minutes of exposure to calling song. The frequencies of trials displaying phonotaxis, non-phonotactic walking and inactivity in the two conditions can be seen in Table 5.1.

The probability of remaining inactive was not significantly affected by the experimental condition (0.54 in the light, 0.50 in the dark, $\chi^2 = 0.10$, $p > 0.05$). In the trials where walking occurred the probability of a track being classed as taxis based on where it reached the wall was 0.91 in the light and 0.67 in the dark, which again is not significant (Fisher exact probability test (two tailed), $p = 0.202$).

Only four crickets walked in both the light and the dark conditions. The paths described are shown in Figure 5.7, with each cricket contributing one or more tracks to each diagram. The tracks collected in the dark condition (right-hand side) initially
### Table 5.1:

<table>
<thead>
<tr>
<th></th>
<th>Light</th>
<th>Dark</th>
<th>Total</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Phonotaxis</td>
<td>10</td>
<td>12</td>
<td>22</td>
<td>(37 %)</td>
</tr>
<tr>
<td>Non-taxis</td>
<td>1</td>
<td>6</td>
<td>7</td>
<td>(12 %)</td>
</tr>
<tr>
<td>Inactive</td>
<td>13</td>
<td>18</td>
<td>31</td>
<td>(51 %)</td>
</tr>
<tr>
<td>Total</td>
<td>24</td>
<td>36</td>
<td>60</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.1: Table to show proportion of trials which led to phonotaxis, walking not classified as phonotaxis or termination due to inactivity. Note different totals of light and dark tests.

appear greatly different in character to those from the light condition. However the apparent difference is due to only two tracks from one animal. When these are removed, as in Figure 5.8 the remaining tracks look much more similar in straightness to those obtained in the light.

![Figure 5.7: Left — Paths for crickets walking in the illuminated arena. Right — Paths for the same crickets walking under infra-red light. N = 4 crickets, some contribute more than one track to each plot.](image)

Because the normalised mean vector is based only on movements relative to the speaker it is possible to combine multiple trials for a single animal, and produce a vector representative of its overall activity. This was done for the four crickets of Figure 5.7. Similarly, a combined equivalent of the tracktime measurement was created by dividing the sum of all $\text{min\_times}$ over the sum of all $\text{actual\_times}$ (see Section 5.2.8) for a cricket. This allowed the D statistic to be calculated for the two conditions. The values obtained are shown in Figure 5.9, and include the looped tracks that were re-
moved in Figure 5.8.

![Figure 5.8: I-R arena tracks with two unusual tracks removed.](image)

While the D statistic is higher in the light than in the dark (means 0.329 and 0.237 respectively) the low number of crickets that walked in both conditions mean it is not useful to try and interpret the scores. To employ the Wilcoxon Matched Pairs Test requires a minimum of six pairs of scores, and significance at the 5 % level is only obtained if all scores change in the same direction.

![Figure 5.9: D statistic for light and I-R tracks, including the unusual looped tracks.](image)

5.5 Discussion

5.5.1 Phonotaxis Without a Striped Grating

The original intention was that the blank arena would provide no input to the cricket’s optomotor system, due to a lack of vertical contrast edges. It was predicted that this
would result in phonotaxis with large oscillations in the heading (up to 60°) as seen in experiments with *G. campestris* on a walking compensator [Schmitz et al., 1982].

Although there were some occasions when a cricket turned away from the speakers by as much as sixty degrees the deviations observed were generally not so large. It is possible that the ridged pattern of the acoustic foam and the corners of the sound-proof box provided enough structure for a significant amount of visual stabilisation. Weber and colleagues observed tracking with deviations of only six degrees when they illuminated the room containing their walking compensator [Weber et al., 1981]. Based on their description of the contents of the room, the amount of visual structure visible to their crickets may have been rather similar to that in the current experiments.

The directness of tracking displayed by a few crickets in the dark with infra-red illumination is surprising (Figure 5.8). The cricket that produced large looping paths in the dark (Figure 5.7, right) suggests that the removal of visual cues might have a more pronounced effect on some crickets than others. As Table 5.1 shows, a larger proportion of crickets produced walking paths that were classed as non-phonotactic in the dark. It may be that these insects were influenced to walk by the presence of the calling song pattern, but just performed particularly badly at localising it in the absence of visual cues. As stated in Section 5.4 this effect has not been shown to be significant, but it is the closest that the current experiments have come to showing the kind of effect that Weber observed on the walking compensator.

### 5.5.2 Phonotaxis with Landmarks and Gratings

It was hypothesised that the introduction of black stripes to the arena (to form a grating) would provide high contrast input to the crickets' optomotor responses, making their courses straighter and more direct. In fact the introduction of a grating, of either solid black stripes or graded stripes, seemed to distract the crickets when they neared the speaker. This resulted in mean vector directions that were less direct. The presence of the black stripes resulted in a (non-significant) increase in the D statistic, but this was mainly due to the crickets tending to run faster. Overall the grating causes too many interdependent effects to say anything conclusive about how the contribution of the optomotor response might be affected.

It was already known from previously published work that black shapes would probably be attractive to the crickets (reviewed in Section 2.5.3) and this was confirmed by some preliminary experiments described in Appendix B, Section B.1. However,
some further experiments (Section B.2) seemed to show that in most cases the crickets would track the calling song and ignore a deliberately placed distractor. A grating made of small rectangles similar to the one used as an attractive target equally seemed to produce no distracting effect, and no course stabilising effect either (results not shown).

It was suggested (T. Collett, personal discussion) that the cricket’s optomotor system might be most responsive to movement in the upper part of the visual field as it is for some other arthropods like crabs [Nalbach and Nalbach, 1987]. The stripes used in Section 5.3.2 were introduced to match the horizontal frequency and vertical extent of the gratings used in the trackball experiments, which were known to be effective. It was only when these gratings were used that the tendency of the crickets to deviate towards the dark stripes during phonotaxis was apparent. Crickets may have preferences for different shapes of target as shown for *Acheta domesticus* [Atkins et al., 1987]. *A. domesticus* were found to prefer horizontally orientated shapes subtending a horizontal angle of 30° or more. Seemingly the *G. bimaculatus* specimens used here preferred the larger stripes.

It does not appear that the crickets in the current experiments exhibited choice behaviour. The calling song and the attractive stripes of the grating both affected their paths, which approached the correct quadrant of the arena, whilst displaying a bias to a nearby stripe.

### 5.6 Summary

It was not possible to demonstrate a significant change in the directness of the crickets’ paths between the different visual conditions tested. Whilst this result is inconclusive it tends to support the idea of an adaptive efference copy, since this would allow the cricket to compensate for the increased optomotor response in a patterned environment. The insignificant changes seen in the cricket’s paths appear to be responses to individual stripes of the grating patterns. The crickets did not display the kind of behaviour that would be predicted from summation at the motor output or a fixed efference copy, that is quicker turning in the dark, or slower turning in the presence of stripes.
Chapter 6

Robotic Model Specification

6.1 Introduction

Following the arena and trackball experiments it was concluded that the most likely mechanism to explain the crickets’ behaviour is a form of efference copy. Summation at the motor output also could not be ruled out; it was still possible that some feature of the dynamics of the auditory or optomotor pathways might enable them to operate successfully in parallel. This chapter describes the design and implementation of a robotic model for evaluating these mechanisms in physical sound-localisation tasks equivalent to the cricket experiments.

Biological relevance was of central concern in the modelling process, which meant constructing the model with reference to the known anatomy and physiology of the cricket. Of course, some parts of the system had more biological data to constrain their construction than others. This, together with limitations of the chosen robotic platform, meant that careful consideration had to be given to the level of description, degree of detail and specificity that could be achieved for each component. Section 6.2 sets out the principles that were used to inform this process. Section 6.3 describes how the forward model used in the robot controller was inspired by findings from studies of the insect mushroom bodies. Finally, sections 6.4 and 6.5 describe the technical details of the robot and its controller.

The finished robot’s control program incorporates three main components. The first, the auditory system, steers the robot reactively towards the source of a cricket calling song. It is based on the data on the auditory neurons that was reviewed in Chapter 3, the previous phonotaxis robot of [Reeve and Webb, 2003], and data from studies of rapid steering in phonotaxis [Hedwig and Poulet, 2004]. The second, the
optomotor system, provides a true optomotor following reflex that rotates the robot in the direction of the dominant direction of optical flow. This is based on the architecture of the robot controller described in [Huber and Bülthoff, 1998], the visual chip of [Harrison and Koch, 2000], and is tuned using new data captured using the trackball system of Chapter 4. These two systems function as independent behavioural modules and can be instantiated independently. Existing models of phonotaxis and optomotor following were relatively mature, having undergone many cycles of refinement and testing. Therefore it was not necessary to change them fundamentally for the new robot. The most important refinement was to match their dynamic properties to those observed on the trackball.

The third component of the robot is the co-ordination mechanism. This may receive input from both of the other two systems, and produces a continuous prediction of the optomotor re-afference. This is accomplished in the model by a recurrent neural network of the liquid state machine type [Maass and Natschlager, 2002]. The co-ordination mechanism can be configured in different ways so that the predictive component learns “off-line” or while incorporated into the optomotor loop. The output can be trained to predict both the magnitude and direction of the response or the magnitude alone, as suggested in Section 4.5.4.3.

### 6.2 Approach

#### 6.2.1 Physical Embodiment

From the outset it was intended that the sensory co-ordination problem should be investigated using a robotic model. The decision to build a model into a physical robot ensures that any proposed controller has to deal with real-world sensory signals. In contrast, when working completely in simulation, sensory inputs may not be modelled in full detail. The danger is that an incomplete or idealised model of the propagation of sensory signals may simplify the task of the simulated agent so that it works in the simulation but would fail in real life.

A cricket’s auditory input during a phonotaxis experiment is a combination of sound travelling directly from a singing cricket or speaker, together with echoes that have reflected from different surfaces with properties causing different levels of attenuation. These factors would make realistic computer modelling of the sensory environment difficult. Similarly, the cricket visual system has to cope with shadows, re-
reflections, fluctuating light levels and moving objects. Whilst techniques for simulating these visual phenomena are better established, doing so would mean the environmental portion of a simulation becoming more complex and time-consuming to build than the current model of the insect.

Limitations of Robotic Hardware

A major limitation of the robotic approach is the type of hardware available, and the effect this has on the interaction of the model system with the environment. This applies to both sensing and action. Electronic sensors are never perfect analogues for the biological sensors they represent, and the mechanics available for robot locomotion and manipulation are likely to be particularly poor representations of biological targets. The most obvious difference for the robot described in this chapter is the use of wheels for locomotion instead of legs. Whilst hexapod walking robots do exist the problem of co-ordinating the joints to produce reasonable walking is an active research area in itself. The most successful schemes devised for control do not necessarily use the legs in ways that make them any more relevant to cricket walking than wheels. In addition, existing hexapod robots are much larger than real insects, meaning that the testing of behaviour has to be done on a correspondingly large scale. This potentially makes comparison with the real insects difficult. For this study, in which the temporal scaling of the optomotor system is important, the ability to translate and turn at speeds similar to real crickets is a key requirement of the robotic platform. This makes the use of a wheeled miniature robot a reasonable choice, if not an ideal one.

In order to justify abstracting away legged locomotion it is necessary to examine the limitations of the task that the crickets and the robot have to perform. Legs and wheels are clearly not equivalent, but within the constraints of “ordinary” locomotion on a flat floor (that is to say the cricket does not move by jumping, or side-stepping) it is perhaps reasonable to say that the wheels can achieve for the robot the kinds of movements that the cricket performs with its legs. More generally, if the limits of the equivalence between a hardware component and its biological target can be defined in this way it may be possible to justify the use of that component, so long as the limitations are respected in the interpretation of the results.
Specification of Complete Sensorimotor Pathways

Commitment to a robotic implementation imposes the requirement that the functioning of entire sensorimotor pathways must be specified. Sensory systems cannot function in isolation, but must ultimately result in motor commands that drive the robot around. In the cricket it is generally the case that sensory systems are relatively well understood, while the pathways that connect them to the thoracic and abdominal ganglia are not well specified. The specification of the motor system was therefore one area of the controller for which a high level of abstraction was employed.

Real-Time Constraints

An additional constraint was imposed by the requirement that the robotic model should run in real-time. The computational resources available to the robot had to be divided among the components of its controller in such a way that each one was simulated in sufficient detail without creating delays that would affect the correct functioning of the model. This led to trade-offs between the level of description of different components.

6.2.2 Description at Neuronal Level and Above

The computer program described in this chapter represents the functioning of parts of the cricket brain. Brain function can be examined at multiple scales, from the interactions of whole regions, through the processing of information within networks of identified neurons, down to the synapses, ion channels, and molecules that underlie the neurons' active properties. In choosing a level at which to base the current model, the first consideration was the level that could reasonably be achieved for the co-ordination mechanism that was the focus of the exercise.

The experimental techniques used in this project did not involve any intervention at the neuronal level and therefore could not help distinguish between alternative hypotheses at the neuronal level. In previous chapters different hypotheses have been illustrated using block diagrams, which can be understood to describe mechanisms in an algorithmic sense. These algorithms consist of filters and arithmetic operations that can be implemented directly in a computer program. In doing so the implication is that certain neurons in the cricket brain act together to compute these filtering or arithmetic operations, without it being known which ones.

It is not the case that information is completely lacking about the neurons involved in sensory integration in crickets. Rather, a little information has been gathered about
many different neurons. This information provides a basis for speculation about the role of certain ones, but any such involvement would have to be demonstrated by electrophysiology. In this situation it seems most reasonable to implement the co-ordination mechanism at an algorithmic level, whilst making clear which identified neurons might be involved and might therefore be suitable targets for future laboratory experiments.

In other areas of the model the neuronal basis for certain processes is better known. This is particularly the case for the auditory system. Where models of neural circuitry are better established and have been shown to produce a good match to physiological data then the inclusion of neural detail can be supported. Even though the outputs of the sensory systems are ultimately to be used as inputs to an algorithmic-level process, if a neural model is the most accurate way of arriving at that input then it should be preferred over some more abstract alternative. The same reasoning can be applied to the inclusion of lower-level detail. If details such as spike timing, dynamic properties of the membrane potential, or synaptic facilitation are considered to be essential to a process then they should be included.

In conclusion, the level of description need not be the same throughout the model, and the chosen level should depend largely on the amount of physiological data available. The common framework for the model may be considered to work at an algorithmic level, but a neural model may provide most appropriate description of certain sub-systems.

6.2.3 Specificity

In the field of auditory processing the cricket is used as a standard organism for the study of pattern recognition and localisation. A large literature on cricket phonotaxis and the associated neurophysiology has grown up because of this, including some information on closely related bush-crickets. Study of non-auditory brain function in insects is mainly conducted in other species. Large flies have been widely used for studies of vision in the past and continue to be used, while the genetic tools available in *Drosophila melanogaster* make it the choice for a wide variety of investigations. Certain systems in cockroaches, locusts, bees and stick insects are also intensely studied.

During the design and construction of the model described throughout the rest of this chapter the cricket literature was drawn on as much as possible. However it was
unavoidable that for certain areas the information from crickets was incomplete and it was necessary to draw on evidence from other species. In most cases it was possible to compare data from several different organisms, in order to assess the likelihood of certain structures or mechanisms being general across many insects.

6.3 Forward Model Based on the Mushroom Bodies

A predictive component such as the one being proposed for the cricket robot must take a signal that indicates the output activity of the auditory system, transforming it into a prediction of the optomotor system's output. The optomotor system's response can vary over time, according to the spatial frequencies present in the visual scene, and the level of contrast provided by that structure. In addition, the exact movements carried out by a cricket in response to sound may vary over time due to injury of the body, fatigue or adaptation. The proposed component should be able to cope with these changes by being plastic. Essentially it has to learn to compute a mapping between temporal functions.

From the literature reviewed in Chapter 3, Section 3.5, it is apparent that several of the features that would need to exist for such a computation are present in the mushroom bodies. It is known that the necessary sensory modalities are represented in the responses of extrinsic neurons at both the calyces and the lobes. There are pathways that mediate reinforcement signals to control plasticity. There are recurrent networks that permit temporally extended responses to acute sensory stimulation.

Temporally Extended Output

As was shown on the trackball, the optomotor system has a dynamic response that attenuates high frequencies (Section 4.3.3). The auditory signals ascending to the brain, on the other hand, copy the short pulses of the calling song with accuracy [Wohlers and Huber, 1982]. This means that a mechanism implementing the forward model must have the ability to generate output that lasts beyond the duration of its input. This has been demonstrated in the extrinsic neurons of both the cricket and the cockroach mushroom bodies [Schildberger, 1984a], [Li and Strausfeld, 1997]. In theory it would be possible to create sustained activity up to a finite limit by using an array of delay lines of different lengths. Mobbs touches upon this idea in his anatomical review paper [Mobbs, 1982], but gives an estimate of the delay that might be imposed in
the Kenyon cell axons as 3.5 ms. On its own, even coupled with synaptic delays, this is not sufficient to generate the kinds of extended activity seen in the extrinsic neurons. The sustained activity observed is most likely to be due to reverberating activity in recurrently connected neurons.

Learning Signals

Neurons and neurotransmitters that substitute for reinforcement in associative learning tasks have been identified [Giurfa, 2006]. The learning of a prediction in a forward model can be seen as a type of operant conditioning. The actions of the insect lead to sensory input and an output from the predictive component. If the two signals do not match then the discrepancy represents an error. If this error is propagated back to the predictive component as a reinforcing signal then this can guide the tuning of the output.

Heisenberg 2003 Model of Associative Learning

The existing data from olfactory conditioning experiments has led to a simple speculative model about the learning processes taking place [Heisenberg, 2003]. Figure 6.1 shows a sketch of the model. Projection neurons that carry information about the presence of a single primary odour quality (POQ) synapse onto a much larger number of Kenyon cells. Each Kenyon cell responds to co-incident activation of a small number of primary odour qualities. This effectively results in a sparsening of the code from the PNs.

Lobe extrinsic neurons cause reflex actions when they fire. Initially their synaptic connections are quiescent and they do not fire to any pattern. However the co-incident activation of a reinforcing signal with presentation of an odour causes potentiation of the KC–extrinsic neuron synapses for the excited Kenyon cells.

In order to change Heisenberg’s model into a predictive component of the type required certain changes have to be made. Firstly Heisenberg’s model takes no account of dynamic behaviour. This is required in the predictive component, so recurrent connections between the Kenyon cells must be added. Secondly, the nature of the outputs of Heisenberg’s model are not specified. It is implied that they act as triggers which, once activated by presentation of a learnt odour, cause a behavioural sequence to be executed. In the predictive model the output must be able to approximate a continuous function. This could be achieved through rate coding in a single neuron, or through the
Figure 6.1: Odours are represented as a population code in the projection neurons (PNs). Each PN reports the presence of a single “Primary Odour Quality” (POQ). Each Kenyon Cell (KC) receives input from a small number of PNs, making it selective for a particular combination of POQs. Extrinsic neurons (XN) receive synapses from all KCs. Initially the KC–XN synapses have zero strength. However the arrival of a targeted modulatory signal reporting reward (apple) or punishment (shock) will potentiate the synapses of the currently active Kenyon cells onto the extrinsic neuron mediating a conditioned response. Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Neuroscience, Mushroom Body Memoir, Martin Heisenberg, Volume 4, Issue 4, Pages 266–275, copyright 2003.
activity of a population of neurons. Thirdly, the reinforcement signal may need to be different. In learning a function the output of the prediction may at times be too high and at times too low. It may be necessary to differentiate between these possibilities, or even transmit the magnitude of the error.

**Liquid State Machines**

Liquid state machines (LSMs) are a type of neural network which have been proposed as a theory about how recurrent circuits found in the brain, particularly mammalian cortex, might be used for learning functions of time-varying input [Maass and Natschläger, 2002]. More recently it has been speculated that the liquid state machine concept could apply to other neural structures in which recurrence seems to play an important rôle, including the cerebellar cortex [Yamazaki and Tanaka, 2007] and the locust antennal lobe [Rabinovich et al., 2008].

The essential idea of the LSM is that a population of recurrently inter-connected spiking neurons, known as a liquid pool, may act as a ‘fading memory’. A brief input will cause a pattern of activity to spread through the network, lasting beyond the stimulus duration and dying out slowly over time. The circuit passes through a series of instantaneous states, defined by the spatial patterns of spiking. Similarly, a continuous input produces a continuous series of changes, so that the state at a given instant reflects not only the current input, but the inputs at previous times. In order for the liquid state machine to be able to reliably compute a function of its inputs these patterns of activity must satisfy the *separation property*, meaning that different inputs must produce distinct spatio-temporal patterns of spiking.

[Maass and Natschläger, 2002] shows that these properties arise for a detailed model of the connectivity of rat somatosensory cortex. The model circuitry was constructed by randomly assigning synaptic properties based on experimentally-determined distributions. A very similar circuitry is used in the robot discussed later in this chapter, the details of which are provided in Appendix E.

Within the recurrent pool of a liquid state machine there is no task-specific circuitry to implement a desired function. That is achieved by a second neural component referred to as a readout. A readout receives inputs from all of the neurons of the recurrent pool, and computes outputs based on the spatial patterns of spiking. Readouts may be memoryless, computing their output only on the basis of the pattern at each instant, although this is not essential. General function approximators such as multi-
layer sigmoidal neural networks may be used, or even linear units such as perceptrons [Haykin, 1999]; by producing a non-linear mapping into a high-dimensional space the recurrent circuit acts as a kernel, enabling linear units to produce a reasonable approximation to non-linear functions [Maass et al., 2004]. Several readouts may take their input from the same recurrent circuit and produce different functions of the input in parallel.

Learning in liquid state machines is generally achieved by supervised modification of the weights between the recurrent pool and the readouts. The learning rule used depends on the type of the readout, for example backpropagation for sigmoidal multilayer networks, the delta rule for perceptrons [Haykin, 1999]. Training can be done off-line by using the recurrent network to generate and save the sequence of spatial patterns for a given time-varying input. Then the learning rule can be used to associate these patterns with the desired output values.

Correspondence between Liquid State Machines and Mushroom Bodies

A liquid state machine is used as the predictive component in the robot of this study, partly on the basis of a speculative correspondence with the mushroom bodies: Firstly, the mushroom bodies are thought to be a recurrent structure, with a mixture of excitatory and inhibitory connections. Secondly, the connection between sensory neurons synapsing onto the calyces represents a mapping into a high-dimensional space (the Kenyon cell population). Thirdly there are extrinsic neurons at the lobes which appear to receive input from a large proportion of the Kenyon cells, in a manner reminiscent of the sampling of the recurrent pool of a liquid state machine by its readouts.

There are clearly limits to what is known about the mushroom bodies that mean it is not currently possible to assess the correctness of this analogy. The most obvious difference between the cortex-like circuitry used by [Maass and Natschlager, 2002] and the mushroom bodies is that the cortex model is homogeneous and the mushroom bodies are highly structured. For instance, in the cortical circuit the inhibitory and excitatory connections are distributed randomly throughout a column, whereas in the mushroom bodies the excitatory connections are expected to occur mainly between Kenyon cells within the pedunculi, and the inhibitory connections are expected to be formed by the extrinsic neurons that run between the lobes and the calyces. This does not rule out the possibility that the mushroom body circuit could have the necessary
properties (sustained non-chaotic activity, the separation property), but without a detailed model it is not possible to know.

Data available from the locust shows the activity of Kenyon cells in response to olfactory stimuli [Perez-Orive et al., 2002] [Broome et al., 2006]. The activity obtained in the Kenyon cells is both temporally and spatially sparse, which is rather different from the activity seen in the cortical model. Again, this does not prove that the mushroom bodies cannot act as a suitable recurrent circuit; the cortical model is only one implementation shown to have the necessary properties. However, it is a notable difference between the mushroom bodies and the circuitry implemented on the robot.

Evidence for the presence of learning signals and plasticity at the lobes of the mushroom bodies is strong, but the exact relationships between Kenyon cell activity, extrinsic neuron activity and reinforcing signals is not known. The learning rule adopted for training the liquid state machine in the robot is the Widrow-Hoff rule (see Section 6.5.6) which is an idealised Hebbian learning rule. In reality, multiple mechanisms are likely to be involved in plasticity in the lobes of the mushroom bodies, and the use of the chosen rule is a convenience in the absence of knowledge that would allow a serious attempt to model them. However, there is no reason in principle why a more realistic learning rule could not be used to train an LSM readout.

### 6.4 Robot — System Overview

A Khepera II (K-Team, Switzerland) was used as the mobile base. It is a miniature robot platform 55 mm in diameter, equipped with two drive wheels and infra-red range sensors. Additional sensors can be stacked on top. For the cricket model a custom-made auditory sensor and a camera were added, as shown in Figure 6.2. Internally a Khepera II has a 25 MHz Motorola 68331 CPU and 512 KB of RAM. The processor and memory capacity were not sufficient to run neural simulations of the desired size in real time, so the majority of the control software ran on a more powerful desktop computer. It had two AMD K7 processors and a gigabyte of RAM, although the principal constraint on the size of the simulation was disk I/O speed, due to the large volume of data being recorded. During testing of the robot it was found that a reasonable limit on the size of the simulation was 150 neurons and 600 synapses.

A minimal program running on board the robot was responsible for sending sensor readings to the desktop computer and receiving motor commands through a serial cable. It also provided rudimentary collision avoidance.
Figure 6.2: Photograph of the Khepera II robot showing the arrangement of the camera and auditory hardware.
6.4.1 Auditory Board Details

The auditory hardware was based closely on that described in [Lund et al., 1997] but it was modified in order to reduce the electronic circuit down to a single board which would fit on the robot under the camera. It models the ear as a pressure difference receiver, although of a simpler sort than the real cricket ear system described in Section 3.2.1. An equivalent physical system would consist of a membrane at the end of a tube which acts as an acoustic wave-guide (Figure 6.3). Sound reaches the membrane both directly from the outside and by travelling along the tube. The output of the system is the amplitude of the membrane oscillations, which depends on the relative phase of the pressure wave at the two sides.

Figure 6.3: Pressure difference receiver as implemented by the auditory circuit. The membrane at the right-hand side moves in response to the difference in pressure between the inside and the outside of the wave-guide.

In Figure 6.3 the tube is positioned so that its open end is facing away from the sound source. In this situation sound reaching the inside of the membrane is delayed by the time it takes to travel along the outside then the inside of the tube; its delay relative to the sound at the external surface is greater than for any other orientation of the tube. When the model ear is in this position it is intended to correspond to a cricket ear with a calling-song source at 90° to its midline, ipsilateral to the tympanal membrane under consideration. In order for the receiver to produce a maximal response at this angle the tube should be made such a length that the wave reaching the inner surface of the membrane is delayed by 180° relative to the outside. This happens for a tube length of one quarter of the wavelength of the sound plus any number of half-wavelengths. At 4.7 kHz the minimal tube length would be \(343 \text{ m s}^{-1}/(4 \times 4700 \text{ s}^{-1}) = 0.018 \text{ m}\). Now, if the sound source is moved around the model ear so that it points at the opening of the wave guide, the sound reaching both sides of the membrane will be exactly in phase and cancel out. This corresponds to presentation of sound from he cricket's
contralateral side. For intermediate angular positions of the sound source the waves no longer sum or cancel perfectly, creating a cardioid directional response for an ideal point sound source.

It is not possible to build a true pressure difference receiver with conventional miniature microphones. Instead the summation of the signals at the membrane is simulated in hardware. Two microphones are used on the robot, placed 18 mm apart. Conceptually the circuit simulates two pressure difference receivers of the type described, with each microphone acting as both a tympanum for one receiver and a "tube opening" for the other. With the 18 mm separation the phase delays due to external time of arrival work as described for the physical system. The quarter-wavelength delays within the tubes are simulated by electronic delay lines. The delayed signal from each microphone is subtracted from the unaltered signal from the other, and the root mean square of the resulting signals are fed to A/D converters on the robot (Figure 6.4). The circuit diagram for the board is shown in Appendix D.

![Circuit Diagram](image)

**Figure 6.4:** Block diagram of the auditory circuit.

Figure 6.5 shows the directionality of the ears circuit found by rotating the robot on the spot in front of a speaker broadcasting a constant 4.7 kHz tone, sampling from both ears at 22.5° intervals. This was done at several volumes, as it was noted that the symmetry of the two ears varied in a volume dependent way. The left-hand plot was obtained using a relatively low volume and is very symmetrical, with the two curves crossing approximately at 0 and 180° and a clear difference between the two ears at
small speaker angles. The right-hand plot was obtained at a higher volume. Here the left ear output can be seen to dominate, and this causes a shift in the crossing points of the two curves to approximately 20° to the right. The cause of the asymmetry is likely to be mismatches in paired analogue components of the circuit board, for example the amplifiers or the microphones themselves. It is not detrimental to phonotaxis, because the behaviour is reactive. Assuming the robot arrives at a bearing of 20° to the speaker it will continue on this wrong heading for a short while. However, as soon as the movement brings the speaker into a sufficiently lateral position turning to the correct side will be initiated.

![Directionality of the ears circuit at two different volumes.](image)

**Figure 6.5:** Directionality of the ears circuit at two different volumes. 0° corresponds to speaker in front of robot. 90° corresponds to speaker to left of robot. Solid lines, left ear. Dashed lines, right ear.

### 6.4.2 Camera Details

The purpose of the camera was to allow the detection of optical flow in the horizontal direction. Because the field of view of the cricket is very wide in the horizontal plane, perhaps 150° either side of the mid-line [Honegger and Campan, 1989], a camera with an equivalent field of view was used. The camera (K2D, K-Team, Switzerland) was supplied so that the axis of the camera was vertical, pointing at a small hemispherical mirror. The whole horizon was reflected in the mirror and projected into the camera’s field of view. Output from the camera was an analogue composite video signal, which was digitised at 25 frames per second by a frame-grabber (Matrox CronosPlus) installed in the off-board computer. This produced colour images at a resolution of 320 by 240 pixels. The camera performed exposure compensation in hardware although
relatively homogeneous lighting in the experimental arena meant that this probably did not alter much during experiments.

6.4.3 Motor System Details

Motion of the robot was controlled by using the API of the Khepera’s built-in operating system to set the speeds of the two wheels independently. These could be adjusted in steps of 1 rad/s, corresponding to a translation speed change of 0.8 cm/s.

By running the two wheels at equal speeds in opposite directions the robot could be rotated in place. This was used to test the visual system in a way that was comparable to the trackball experiments. The distance between the wheels was 5.3 cm, meaning that the robot’s rotation about its central axis could be adjusted in steps of 0.3 rad/sec.

Under normal operation translation as well as rotation was permitted. The basic translation speed was set at 4 cm/s. Only the rotational component of motion was under the control of the neural model. Rotation was always imposed by speeding up one wheel and slowing the other by the same amount, ensuring that translation speed remained constant. Whilst it is clearly not the case that crickets walk at a constant rate all the time (see Appendix C) it is not known what causes a cricket to halt during phonotaxis. Pauses and bouts do not seem to be correlated with the chirps of the calling song, for example, so there seemed little to be gained by including them.

6.4.4 Real-Time Control

The Khepera robot’s operating system allowed periodic sensor polling using built-in API calls. It was capable of sampling the ears reliably at speeds of 1 kHz, allowing good preservation of the shape of individual syllables of a synthetic cricket calling song. This could be done without overloading the robot’s 57 kbps serial line: the ear values were 10 bits each, which could be packaged into three bytes resulting in an average transfer rate of 24 kbps.

This 1 ms temporal resolution was much better than the resolution available on the desktop computer (10 ms). The Khepera was therefore made to act as the real-time clock for the system. At the start of an experiment a message passed to the robot would instruct it to run for a fixed number of cycles. It would start sending ear readings at 1 ms intervals, while the off-board computer would repeatedly wait for inputs on its serial port. Even if the desktop computer started to accumulate data in its serial port buffer the robot would continue to send readings at the same rate, ensuring that the
values used by the control program had been sampled at the correct interval.

1 ms was judged to be a reasonable time step for the neural simulation, so the
arrival of a new ear reading was used as the trigger to update it. The optomotor sub-
system of the robot operated on a longer time-step, constrained by the frame rate of the
video signal. Updates were performed asynchronously in a separate thread of control
running in parallel to the ear-reading process. This thread repeatedly waited for a frame
to be available from the frame-grabber, before carrying out a series of steps to calculate
optical flow as described in Section 6.5.4. On calculation of a new optical flow value
a flag was set in shared memory to warn the main control process to update its optical
flow input on the next time step.

6.5 Robot — Neural Controller

6.5.1 Overview

Control of the robot was accomplished using a neural network which is most eas-
ily understood as three sub-systems: the auditory, optomotor and predictive systems.
The auditory system can be operated independently of the other two systems and
will drive the robot towards the source of a cricket calling song. Internally it uses
a neural network that is changed little in functional terms from the one described in
[Reeve and Webb, 2003]. It models the identified cricket auditory neurons and by do-
ing so replicates the pattern selectivity of females.

The optomotor system provides an optomotor following reflex; if it is active it will
rotate the robot in the direction of the predominant optical flow direction across the
visual field. It is based closely on two previous models of optomotor behaviour in flies
[Huber and Bülthoff, 1998] [Harrison and Koch, 2000].

The predictive system can receive input from either or both of the other two sys-
tems. Its function is to produce an on-going prediction of the output of the optomotor
system and subtract this quantity from the optomotor reflex pathway, in order to allow
the robot to respond only to externally generated optical flow. It is based on a liquid
state machine (LSM).

6.5.2 Neuron Models

The controller's networks are built from spiking neurons, which explicitly model the
dynamics of the membrane potential and the timing of individual spikes. Two types
Figure 6.6: Architecture of the neural controller. Solid lines—Signals of the reflex pathways. Dotted lines—Signals of the predictive sub-system. Open circle—See Section 6.5.6 for details of combination for different variants of the predictive model.

Chapter 6: Robotic Model Specification
6.5. Robot — Neural Controller

6.5.2.1 Leaky Integrate-and-Fire Neuron

The predictive system uses leaky integrate-and-fire neurons [Dayan and Abbott, 2001]. Neurons are treated as a single compartment with capacitance $C$ and resistance $R_m$. The resting membrane potential, $v_{\text{rest}}$, and the threshold voltage at which a spike occurs $v_{\text{thresh}}$ are also parameters. The evolution of the membrane potential of the cell, $v_m$, is determined by the equation

$$C \frac{dv_m(t)}{dt} = -\frac{1}{R_m} (v_m(t) - v_{\text{rest}}) + I(t)$$

where $I(t)$ is the current contributed by synapses plus any externally injected current.

Spikes are said to occur when the membrane potential reaches $v_{\text{thresh}}$, at which point it is reset to a new value, $v_{\text{reset}}$, and integration can begin again. The behaviour of the membrane potential during a spike is not modelled; in the context of the model a spike is a signal, which is used to modify the behaviour of model synapses. The integrate-and-fire neurons of the predictive system have refractory periods of few milliseconds. These are enforced by suppressing the spiking mechanism for the required duration after a spike. During this time integration continues as normal.

6.5.2.2 Quadratic (Izhikevich) Model

The auditory system uses a more complex two-variable neuron model [Izhikevich, 2007a] which takes into account the voltage-gated currents that cause spiking and slow currents that re-polarise the neuronal membrane after a spike. This model allows much more realistic membrane dynamics to be simulated at the expense of extra computation time. Because different types of biological neurons have quite different membrane dynamics the model’s parameters have to be adjusted to get the right behaviour. In theory this can be done if one knows certain electrical properties of the cells being simulated. However the model is still a great simplification of the processes taking place in a real neuron, so not all the observed properties can be expected to emerge from correctly matching the four parameters.

As for the leaky integrate-and-fire neuron, the quadratic model assumes a single compartment. Certain parameters, namely $C$, $I(t)$, $v_m$, $v_{\text{rest}}$ have the same meanings. There is also a 'threshold' parameter, $v_{\text{thresh}}$ although this does not denote the membrane potential at which the cell is guaranteed to fire as it does in the previous model.
Additional parameters of the quadratic model are denoted $u, a, b, c, d$ and $k$. Izhikevich gives intuitive explanations of their meaning in [Izhikevich, 2003]. The most important is $u$, which

[...] accounts for the activation of $K^+$ ionic currents and inactivation of $Na^+$ currents, and [it] provides negative feedback to $v$.

The equations governing the evolution of $v_m$ and $u$ are

$$C \frac{dv_m(t)}{dt} = k \cdot (v_m(t) - v_{rest}) \cdot (v_m(t) - v_{thresh}) - u(t) + I(t)$$

and

$$\frac{du(t)}{dt} = a \cdot (b \cdot (v_m(t) - v_{rest}) - u(t))$$

Izhikevich’s model, unlike the leaky integrate-and-fire model, does account partly for the membrane dynamics during a spike. The above equations will produce the upstroke of the spike. Once $v$ has reached a certain level, $v_{peak}$, the model is reset according to the auxiliary equations

$$v_m \leftarrow c$$

and

$$u \leftarrow u + d$$

The condition $v_m \geq v_{peak}$ is also used as the trigger to send a spike signal to the neuron’s synapses.

In theory this more complex type of neuron model could be used in the predictive sub-system. In practice the potentially realistic membrane dynamics gained from using the two-variable model are only useful if they can be matched properly to the cells being modelled and contribute in a meaningful way to the simulated behaviour. In the auditory system intrinsic properties of small numbers of neurons may play a part in the recognition of calling song patterns [Bush and Schul, 2006], and the two-variable model may be adjusted to experiment with these properties. In the mushroom body the dynamics of individual cells and the connectivity between them are much less certain. Given the state of current knowledge about the exact circuitry in this part of the cricket brain there is little point in speculating about the role of fine temporal dynamics of the membrane potential, so the simpler integrate-and-fire model is used.
6.5.2.3 Axons

Neurons of the model may be defined to have an axonal delay. These introduce a lag between the firing of a neuron and the time at which the synapses receive the spike signal. They are implemented using a queue data structure.

6.5.2.4 Linear Readout Unit (ADALINE)

At only one place in the system is a non-spiking neural model used. A linear unit, or ADALINE [Haykin, 1999] is used as the readout of the liquid state machine in the predictive sub-system. A linear unit computes the dot product between its weight vector $w$ and its input vector $x$:

$$ y = w \cdot x $$

The output $y$ can be loosely interpreted as the firing rate of a neuron.

The use of such a model for the readout was not unavoidable, but preliminary attempts to train integrate-and-fire neurons to perform the same function resulted in poor results. Because the coding of mushroom body extrinsic neurons is assumed to be rate-based for the purposes of the model it was considered satisfactory if not ideal to use the less realistic linear unit.

6.5.3 Synapse Models

The behaviour over time of a spiking neuron model is governed by the sum of its input currents

$$ I(t) = \sum_i current_i(t) + I_{ext}(t) $$

where $current_i(t)$ is the contribution of the $i^{th}$ synapse. For conceptual clarity any externally injected current (from an experimenter's electrode) is separated out from the synapses and denoted $I_{ext}$, although in practice it contributes in exactly the same way. Four types of model synapse are used.

6.5.3.1 Conductance-based Synapse

The most biologically realistic type is used in the auditory system. There, the contribution of a synapse in nanoamperes is dependent on its reversal potential, $r_i$ millivolts, and the post-synaptic membrane potential, $v_m$ millivolts, according to

$$ current_i(t) = g_i \cdot S_i(t) \cdot (r_i - v_m(t)) $$
where $g_i$ is the intrinsic synaptic conductance, or weight, in microsiemens and $S_i$ is its level of activation. Setting the reversal potential, $r_i$, to below the resting potential of the post-synaptic cell produces currents that tend to hyperpolarise the cell and hence inhibit it from firing. Setting $r_i$ to a value above the spiking threshold of the post-synaptic cell produces currents that generally depolarise the cell and may elicit spikes.

The variable $S_i$ determines the shape of the post-synaptic potential. When a neuron spikes, $S_i$ is incremented in all the synapses to which it is pre-synaptic. The size of the increase in each one is determined by the parameter $\delta_i$. $S_i$ then decays exponentially with a time constant, $\tau_i$. These dynamics approximate the binding then unbinding of neurotransmitter from post-synaptic receptors.

### 6.5.3.2 Depressing Synapse

Depressing synapses are a type of conductance synapse that can also undergo short term depression, corresponding to the exhaustion of the pre-synaptic vesicle pool. This is modelled by an additional variable $d$:

$$\text{current}_i = g_i \cdot (1 - d_i(t)) \cdot S_i(t) \cdot (r_i - v_m(t))$$

The temporal properties of the depression are determined by parameters $\delta_{\text{dep}}$ and $\tau_{\text{dep}}$. $d$ is initially zero. Presynaptic spikes cause it to be incremented in steps of $\delta_{\text{dep}}$ up to a maximum value of one, when the synapse no longer makes a contribution to the input of the postsynaptic cell. In between spikes it decays exponentially with the time constant $\tau_{\text{dep}}$. This effect means that a single synapse can act as a low-pass filter for pre-synaptic rate. When the rate is low $d$ recovers in between spikes and transmission can go on continuously. As the rate becomes higher the depression variable reaches its maximum, and a reduction in the presynaptic spike rate is required before transmission of spikes can resume.

### 6.5.3.3 Exponential Synapse

The third and simplest type of synapse is modelled as an exponentially decaying current which is independent of the state of the post-synaptic membrane. The peak current ($A$) and the time constant of the decay ($\tau$) are parameters. This type of synapse is used as an output from motor neurons which cannot use the above schemes because, for the purposes of the model, they do not have any post-synaptic targets.
6.5.3.4 Full Dynamic Synapses [Markram et al., 1998]

The final type of synapse is used only between the neurons of the recurrent pool in the predictive model. It incorporates facilitation as well as depression, using the model of [Markram et al., 1998]. As for the simplest synapse described above the post-synaptic potentials are exponentially decaying currents which are independent of the state of the post-synaptic membrane. The amplitude of each potential is determined by three factors. $A$, in nano-amperes, is the inherent weight of the synapse. $R(t) \in [0, 1]$ is initially 1 and represents the fraction of available synaptic efficacy. It is decreased following a spike and thus mediates depression. $u(t) \in [U, 1]$, where $U$ is a base level supplied as a parameter, determines the proportion of synaptic efficacy that is used up by a spike. The arrival of a spike causes an increase in $u$ of $U(1-u)$, thus mediating facilitation whilst limiting the value of $u$ to 1.

Denoting the levels of the various factors just before the $n$th spike by the subscript $n$, the amplitude of the $n$th spike is $A \cdot u_n \cdot R_n$, and the amount of synaptic efficacy remaining after the spike is $R_n - u_n \cdot R_n$. In between spikes $u$ and $R$ decay exponentially to their base values ($U$ and 1 respectively) according to chosen time constants $\tau^{facil}$ and $\tau^{dep}$.

The behaviour of the model is illustrated in Figures 6.7 and 6.8, which show the behaviour of current, $R$ and $u$ when dynamic synapses of different parameterisations are presented with a spike every 500 ms starting from rest. Figure 6.7 shows a synapse with its parameters set to the mean values for excitatory–excitatory connections in the recurrent pool of the predictive component: $A = 30$ nA, $U = 0.5$, $\tau^{dep} = 1100$ ms and $\tau^{facil} = 50$ ms. Figure 6.8 shows the effect of the mean parameter values for excitatory–inhibitory connections: $A = 60$ nA, $U = 0.05$, $\tau^{dep} = 125$ ms and $\tau^{facil} = 1200$ ms. Note the different scales on the ordinate.

With the first parameterisation the relatively long time constant for $R$ means that depression dominates the behaviour. The first pulse of current is the largest, with subsequent pulse heights approaching an asymptote of approximately 8.0 nA, which depends on $U$, $\tau^{dep}$ and the inter-spike interval. Although the parameter mediating facilitation, $u$, is incremented at each spike its short time constant means that by the time the next spike arrives it has decayed back to its base level. This is in contrast to the behaviour of the second synapse, where the relative lengths of the two time constants are reversed. In this case $R$ has a negligible effect on the size of each spike; the full

---

The equations of [Markram et al., 1998] contain a small error which is corrected in [Maass and Markram, 2002]. The corrected version is used here.
Figure 6.7: Behaviour of a dynamic synapse presented with spikes every 500 ms. $A = 30\ \text{nA}, U = 0.5, \tau_{\text{fac}} = 50\ \text{ms}, \tau_{\text{dep}} = 1100\ \text{ms}$. Top to bottom: current, $R$ and $u$.

Figure 6.8: Behaviour of a dynamic synapse presented with spikes every 500 ms. $A = 60\ \text{nA}, U = 0.05, \tau_{\text{fac}} = 1200\ \text{ms}, \tau_{\text{dep}} = 125\ \text{ms}$. Top to bottom: current, $R$ and $u$. 
efficacy of the synapse is always available. Meanwhile $u$ is being incremented each time, resulting in an increase in pulse size up to an asymptotic value of approximately 8.0 nA, which depends on $U$, $\tau^{\text{facil}}$ and the inter-spike interval. In more realistic situations the rate of arrival of spikes would vary over time, causing facilitation and depression to interact.

### 6.5.4 The Optomotor Sub-System

Computation of optical flow begins with the extraction of 360 intensity values, intended to correspond to 360 visual sampling units (VSUs) around the horizon of a cricket’s eye. This may be an underestimate of the average $G. \text{bimaculatus}$ interommatidial angle (Section 3.3). Interommatidial angles of as low as 1° are recorded by Labhart for $G \text{campestris}$, but he was looking specifically at the dorsal rim area of the eye, which may not be typical [Labhart et al., 2001].

The experimental arena around the robot was reflected onto a fixed circular region of the camera image. Exact measurements of the mirror geometry were not known, so the location of the horizon was found empirically by driving the robot forwards and looking for the focus of vertical expansion. The radius of this point from the centre of the mirror’s image was noted, and the circle on which it lay was denoted the horizon. A circular band of the image, four pixels either side of the horizon, was sampled at 0.5° intervals and converted to grey-scale to produce an 8 by 720 array of intensity values. This was then averaged in the vertical direction to produce 720 “ray” values, which were sampled to produce the VSU values.

Subsequent processing of VSU values to detect optical flow was done using a series of filters which provide a good fit to processes known to take place in the optic lobes of flies, without considering the underlying neural circuitry. The exact filters used were based strongly on [Huber and Büthoff, 1998] and [Harrison and Koch, 2000], with the filter parameters adjusted to better fit the cricket data from Chapter 4. The flow of processing is shown schematically in Figure 6.9.

While most of the rest of the robot's controller was implemented using spiking neurons this higher level description of optic lobe function was reasonable for several reasons. Firstly, for assessing the performance of the predictive part of the controller it was only necessary that the optomotor input had the right temporal structure. Whilst theories about the circuit basis of motion detection exist and have been modelled [Higgins et al., 2004], the output generated by this circuit was similar enough
Although the inter-axis angle of visual sampling units will correspond to the inter-ommatidial angle of the eye the acceptance angle of a VSU may be larger, causing some overlap in the points in space sampled by adjacent units. This is simulated in the calculation of VSU intensity values from ray values. Ray values around the axis of each VSU are sampled according to a Gaussian distribution with standard deviation 1.65 rays, or 0.825°.
6.5.4.2 Temporal high-pass filtering in the lamina

The transient nature of the response of lamina monopolar cells [Laughlin, 1987] is simulated using a high-pass filter with $\tau = 58$ ms.

6.5.4.3 Local motion detection in the medulla

Local motion detection between adjacent VSUs is accomplished using the Hassenstein-Reichardt model of elementary motion detection [Hassenstein and Reichardt, 1956] (see also Section 2.8.2). The variant used employs first-order low-pass filters to produce the necessary delays and multiplication to perform the non-linear combinations. Each VSU forms an input to two elementary motion detectors (EMDs, see Figure 2.17), one with each of its neighbours, so that the number of EMDs equals the number of VSUs. The time constant, $\tau$ of the low-pass filters was set to 93 ms. This gave the characteristic curve of Figure 6.10). The curve is a reasonable fit for that of the cricket (Figure 4.9), with maximum turning for a stimulus pattern frequency of approximately 2 Hz, dropping off to 60% of this value at 5.5 Hz.

![Figure 6.10: Characteristic curve for robot optomotor system in response to unidirectional pattern motion](image)

6.5.4.4 Spatial summation in the lobula

The last stage of visual input processing is the summation of the EMD outputs across the visual field. This corresponds to integration in tangential cells of the lobula.
6.5.5 The Auditory Sub-System

The portion of the robot's controller concerned with auditory processing consisted of the neural network and pre-processing steps shown in Figure 6.11. It had two main functions: pattern recognition and enhancement of the ears' directionality.

Figure 6.11: Pre-processing stages and neural network of the auditory subsystem. AN – Ascending Neurons, ON – Omega Neurons, BN – Brain Neurons. Open arrowheads – excitatory synapses, open circles – inhibitory synapses, closed arrowheads – current-based synapses.
6.5.1 Pre-Processing

Output values from the auditory hardware of the robot were digitised with ten-bit precision and passed to the off-board PC. At background noise levels the left and right channels both had roughly constant outputs of consistently different magnitudes. A subtractive bias was introduced in order to equalise the two channels at zero under such conditions. The bias was subsequently increased, by an equal amount on both sides, so that it was sufficient to remove spikes of noise that became a problem when the motor system of the robot was active. The resulting signal was half-wave rectified in order to ignore negative values. Rectified signals were multiplied by fixed gains to produce values in the range 0 to 18640 which, as simulated current inputs, were the right order of magnitude to produce spikes in the receptor neurons.

6.5.2 Spiking Neural Network

All neurons were spiking neurons of the quadratic model type (see Section 6.5.2.2) with the parameters set as follows: \( a = 0.03; b = 8; c = -45; d = 20; k = 1; C = 20; v_{rest} = -55; v_{thresh} = -40; v_{peak} = 25 \). The initial conditions were \( v = -60 \) and \( u = 0 \). \( I_{ext} \) was 0, except in the case of the receptor neurons which only received “injected” input.

The receptor neurons perform the function of the auditory nerve in the cricket. In reality there are around sixty primary nerve fibres in the cricket, but using a single neuron as described here was sufficient to produce the right pattern of firing in the primary auditory interneurons. Apart from the output synapses from the BN cells the synapses were of the conductance-based type (Section 6.5.3.1). Those between the AN and BN neurons incorporated the depression mechanism discussed in Section 6.5.3.2, with \( \tau_{dep} = 10 \) and \( \delta_{dep} = 0.25 \). Parameter values for the different connections are given in Table 6.1. \( \tau \) was 5 ms in all cases. Because the circuit is symmetrical values for the left and right sides are not listed separately.

Four identical excitatory synapses transmitted spiking activity in the receptor neurons to the AN1 and ON1 neurons. ON1L and ON1R were mutually inhibitory, and also inhibited the AN1 neurons on the opposite side. At the onset of sound, spiking would tend to occur sooner in the receptor and ON1 on the side where the sound is loudest. This would inhibit the ON1 and AN1 neuron on the quieter side, tending to prevent spikes from being initiated there, and thus enhance the signal directionality.

At the standard syllable repetition interval the AN1 neurons copy the sound pattern
Table 6.1: Synapse parameters for the auditory sub-system

<table>
<thead>
<tr>
<th>Connection</th>
<th>Type</th>
<th>( r ) (mV)</th>
<th>( \delta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Receptor to AN1</td>
<td>Excitatory</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Receptor to ON1</td>
<td>Excitatory</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>ON1 to opposite ON1</td>
<td>Inhibitory</td>
<td>-90</td>
<td>5</td>
</tr>
<tr>
<td>ON1 to AN1</td>
<td>Inhibitory</td>
<td>-90</td>
<td>3.75</td>
</tr>
<tr>
<td>AN1 to BN</td>
<td>Excitatory</td>
<td>0</td>
<td>3.75</td>
</tr>
</tbody>
</table>

rather faithfully (see Robot Model Results, Figure 7.2). The brain neurons, BN, have a more complex response because the synapses that connect them to the ANs undergo short term depression. This depression, which recovers with a time constant \( \tau \) of 10 ms means that the brain neuron tends to respond only to the start of a syllable, and requires a quiet inter-syllable interval to allow full transmission to recover. To a certain extent this means that the brain neurons will be more active in response to calling songs that contain more syllables per chirp. At very short SRIs, however, the ANs no longer copy the calling song very faithfully, with no discernible gaps in the spiking pattern. At the level of the brain neurons chirps with such short SRIs appear the same as a long continuous syllable, eliciting a response only at the onset.

Output from the auditory sub-system is via two exponential synapses (Section 6.5.3.3). For each spike in the brain neurons the output of these synapses jumps by a fixed amount. In between spikes they decay with a time constant of 5 ms. In order to allow an algorithmic combination with the output of the optomotor system the value of the left synapse is subtracted from that of the right, producing a continuous quantity which is positive for rightward turning and negative for leftward turning. The resulting signal is filtered by a low pass filter with a cut-off frequency of 3 Hz to approximate the dynamics of rapid steering seen on the trackball (Section 2.4.3).

6.5.6 The Predictive Sub-System

The predictive sub-system is a liquid state machine as described in Section 6.3. The recurrent “liquid pool” is based on the same mixture of excitatory / inhibitory neurons and connection probabilities outlined in [Maass and Natschläger, 2002]. Full details are given in Appendix E. As noted in Section 6.3 this assumes a statistically homogeneous structure which, for a model of the mushroom bodies, is a simplification.
Inputs to the recurrent network represent efference copies from either the auditory system or the optomotor system or both. As described in Sections 6.5.4 and 6.5.5 the outputs of the optomotor and auditory sub systems are one-dimensional real-valued variables that reflect the intended rotation of the robot. The liquid pool requires spike trains as input that give the same information. For the auditory circuit this is accomplished by using the spike output of the BN cells. These are transferred to the predictive sub-system via two delay lines (axons) which can be adjusted to take into account some of the lag inherent in the robot’s optomotor hardware. For the optomotor system, which is not internally composed of spiking neurons, the output value is transformed into spike trains by two spiking neurons, one for rightward and one for leftward optical flow. The output of the optomotor system was “injected” as a current input into the appropriate neuron depending on its sign. In order that the neurons could preserve the information about pattern speed the current input was multiplied by a gain factor that ensured that the neurons’ responses would not saturate at the peak output of the EMDs. Each input spike train (two auditory and possibly two optomotor) was connected by excitatory synapses to randomly chosen 30% of the neurons of the recurrent pool.

The readout used for the liquid state machine was a single linear unit (ADALINE) [Haykin, 1999]. At a given instant $t$, if the vector of filter outputs from the recurrent pool is $\mathbf{x}(t)$ and the weight vector of the readout is $\mathbf{w}$, then the prediction of the liquid state machine is the dot product $\mathbf{y}(t) = \mathbf{w} \cdot \mathbf{x}(t)$. If the observed optomotor reafference at time $t$ is $o(t)$ then the difference $e(t) = o(t) - y(t)$ is the error on that time step. Training the readout consists of altering the weight vector in order to minimise $e$ over all $t$. In other words it requires finding a solution to an overdetermined system of inconsistent linear equations $X\mathbf{w} = \mathbf{y} + \mathbf{e} = \mathbf{d}$ that minimises some appropriate function of the error vector $\mathbf{e}$ such as the mean square.

Solving such a system of equations can either be done off-line, using a method for linear regression that produces an optimal fit, or on-line. The Widrow-Hoff Delta Rule [Widrow and Lehr, 1990] is an algorithm that can be used for on-line training. This uses the error at time $t$, $e(t) = o(t) - \mathbf{w} \cdot \mathbf{x}(t)$, to adjust the readout weights according to $\mathbf{w}_{t+1} = \mathbf{w}_t + \mu e(t) \mathbf{x}(t)$ where $\mu$ is the learning rate.

The output of the predictive sub-system was combined with output of the optomotor system at the point indicated by the empty circle in Figure 6.6. The algorithm for combination differed depending on the variant of the predictive sub-system. If the directional version of the predictive model was in use then the readout values took both negative and positive values and the combination was a subtraction:
motor_signal = phono_signal * phono_gain
    + (opto_signal - prediction) * opto_gain;

If the non-directional version of the predictive model was used then the readout was trained to predict only the absolute value of the optomotor re-afference. A half-wave rectification function was applied to the output of the predictive sub-system to eliminate any negative outputs. The resulting signal was combined with the optomotor output according to the following algorithm:

discrepancy = prediction - abs(opto_signal);
current_steering = sign(phono_signal);
motor_signal = phono_signal * phono_gain
    + (opto_signal + discrepancy * sign) * opto_gain;

where \( \text{abs} \) computes the absolute value and \( \text{sign} \) returns 0, 1 or -1 depending on the sign of its argument.

Informally, the non-directional combination determines the discrepancy between the expected and the observed optomotor magnitude, and attempts to minimise the discrepancy by making the robot turn faster or slower towards the sound as required.

6.6 Fitting the Reflex Pathway Dynamics to the Cricket Data

6.6.1 The Auditory Output Filter

In rejecting the chained subsystems hypothesis some consideration was given to the filtering properties of phonotactic steering (Section 4.5.2). It was concluded that a first-order low-pass filter with a cut-off frequency of 3 Hz at the output of the auditory system would provide a good fit to the cricket data from [Hedwig and Poulet, 2005]. Such a filter was therefore implemented on the robot. To check the output for a reasonable fit to the cricket the robot was placed at the centre of the arena of Chapter 5 with two speakers positioned 40° to the left and right of the midline. The motor system was disabled, then the output of the auditory integrator was logged during 30 seconds of split song (see Section 2.7) and 30 seconds of one-sided standard calling song (as described in Section 5.2.2). The resulting output signal of the auditory system was then averaged across the chirp period and plotted in Figure 6.12.
Like the cricket the robot is capable of responding to each pulse of a split song. As described in Section 7.3, the output of the BN neurons is likely to consist of a single spike per chirp. The filter smooths out the spikes, giving a better match to the mean walking pattern seen in the insect.

Calling song from one side only produces a turning tendency towards the speaker, with a peak amplitude of 0.71 and negligible DC shift. Because the split song output is a noisy average of several trials the oscillations seen in the output trace are not all of equal size. However, if the height from the final trough to the final peak is measured a value of 0.55 is obtained, giving a comparable ratio to that seen in the cricket data.

6.6.2 Frequency Response of the Optomotor Pathway

The first stage in tuning the optomotor subsystem had been to alter the time constant of the low-pass filter in the H-R motion detectors to obtain a static characteristic curve with the correct optimum (Figure 6.10). Having done this, the frequency response characteristics of the system were checked. This was done using an equivalent experiment to the cricket tests with sinusoidal input, described in Section 4.3.3. The robot was placed in the centre of the sound-proof arena, with the solid striped pattern of Figure 5.2 fixed to the walls. The robot's controller was disconnected from the motors, and sinusoidal motor commands were injected to cause oscillation in place. The outputs of the optomotor system were recorded.

When testing the crickets with sinusoidal inputs the grating's range of angular po-
sitions had been restricted. For the robot the range of speeds was fixed, to simplify the analysis. The motor speeds were set according to the equations

\[ mot_r = \text{rint}(10\sin(2\pi ft)) \]

and

\[ mot_l = -\text{rint}(10\sin(2\pi ft)) \]

where \( t \) is time in seconds, \( f \) is the oscillation frequency in Hertz and the \( \text{rint} \) function causes rounding to the nearest integer. Rounding is necessary because the robot’s motor system only accepts speed instructions in integer increments. The peak speed of ten units corresponds to an angular velocity of 172°/s or a pattern frequency of 2.86 Hz. Four cycles-worth of output were captured at each of seven speeds: 0.05, 0.1, 0.2, 0.4, 1, 2, and 4 Hz.

The qualitative features of the robot’s optomotor output can be seen in the example traces of Figure 6.13. The upper graphs show the robot’s rotation speed, and the lower ones show the output of the optomotor system from a single trial. Responses of the robot were much less variable than responses of the cricket, so averaging was not necessary.

**Figure 6.13:** Output of robot optomotor system at summation stage in response to sinusoidal inputs at 1 Hz and 2 Hz. Upper traces — left wheel speed (units of 0.3 rad/s). Lower traces — output of summed EMDs (grey lines, arbitrary units) with the dominant component of Fourier transform superimposed (black lines).

Despite the fact that the range of speeds used overlapped with the non-linear range of the characteristic curve the output traces are again dominated by the frequency of
the input signal. Further analysis was therefore carried out by considering this component of the response only. Fourier transforms were calculated for each of the seven output traces, and the amplitudes and phase lags of the fundamental frequencies were extracted. These are presented as a Bode plot in Figure 6.14.

![Figure 6.14: Output of robot optomotor system after summation stage. Left — Amplitude/frequency plot. Right — Phase/frequency plot. Dashed lines & triangles show cricket data, solid lines & circles show robot data.](image)

The Bode plot shows that the model optomotor system itself clearly acts as a low-pass filter before any subsequent integrator is introduced. When compared to the cricket data the cut-off frequency looks higher — around 1 Hz, judging from the amplitude plot. The phase lag curve is similar to that obtained from the cricket. Both curves are dominated by the effects of a fixed time delay. In the case of the robot the delay is inherent to the video capture hardware and computation of the motion signal; it is only fortuitous that the match is a good one for the cricket data. The observed phase data match well with the expected figures for a first-order low pass filter with cut-off frequency 1 Hz plus a fixed delay of 74 ms (not shown).

It is common for models of optomotor following that employ the H-R model to incorporate a separate low-pass filter after the summation of the EMDs. Given the frequency response obtained above it was questionable whether anything would be gained by doing this. The main reason for adding such a filter would be to obtain realistic attenuation of frequencies between 0.3 and 1 Hz. At the summation stage in the current model frequencies below 1 Hz are hardly attenuated at all. However, the slope of the model amplitude/frequency curve above 1 Hz is in good agreement with the cricket data. Adding a separate low-pass filter would make the rate of attenuation in this region too steep. The problem is illustrated by Figure 6.15, in which the experiment has
been repeated with the summed EMD output being filtered by a separate low-pass filter with \( \tau = 499 \) ms. The robot attenuates frequencies between 2 and 4 Hz by up to 20 dB more than the cricket. The phase/frequency curve also looks less satisfactory, with an extra quarter-cycle added to the phase lag at high frequencies. On this basis, coupled with the results of the step response test described in the next section, it was decided to use the optomotor system without any further filtering.

![Figure 6.15: Output of robot optomotor system after summation and low-pass filtering with \( \tau = 499 \) ms. Left — Amplitude/frequency plot. Right — Phase/frequency plot. Dashed lines & triangles show cricket data, solid lines & circles show robot data.](image)

6.6.2.1 Step Response

A step response was obtained for the robot in the same setting used for the sinusoidal input tests described above. The motors were instructed to accelerate the robot as fast as possible from stationary up to a rotation speed of 120° per second. Figure 6.16 shows the output at the summation stage, together with the response of a subsequent low-pass filter with \( \tau = 499 \) ms. Reading off the half-times for the two curves gives a value of around 200 ms for the unfiltered output and 600 ms for the filtered version. Comparing these to the value obtained for the cricket data (Section 4.3.1) shows that the unfiltered output is a better match.

6.7 Setting the Scaling of the Motor Outputs

Before the robot could be tested in an arena setting it was necessary to specify how the outputs of the auditory and optomotor sub-systems should be scaled to produce
6.7. Setting the Scaling of the Motor Outputs

Figure 6.16: Output of robot optomotor system during a step change in rotation occurring at time 0. Solid line — output at summation stage. Note that the value changes only every 40 ms due to the frame rate of the camera. Dashed line — low-pass filter output, \( \tau = 499 \) ms. Dotted lines indicate half maximum output and half-time.

realistic turning speeds. The close-up recordings of Appendix C provided direct measurements of turning speeds during phonotaxis, but the only measurements of steering to optomotor stimuli came from the trackball. The quantitative correspondence between left-right speeds on the trackball and rotation in the arena had to be estimated in order to compare the sizes of the two types of response.

Figure C.6 shows that crickets that paused at 45° to the active speaker turned at approximately 20° per second during the next walking bout (Figure C.7). On the trackball, turning to sound at 45° resulted in a left-right steering speed of approximately 0.3 cm/s (Figure A.4). Assuming a linear relationship between the two types of steering measurement, the following correspondence is established: 1 cm/s L–R trackball speed \( \rightarrow \) 67/s arena rotation speed.

Turning to a moving grating on the trackball resulted in a maximum left-right speed of 1.4 cm/s. Based on the previously obtained ratio (1 cm/s:67/s), a reasonable gain for the optomotor system might be chosen to give a rotation speed of 80–100/s at the most effective pattern frequency. Figure 8B of [Böhm et al., 1991] suggests this is too high, showing a rotation of 35/s when a 1 Hz grating is combined with a calling song from the crickets’ front.

Ideally the correspondence would be established through experiments with a better equivalence than those chosen above. For instance the arena experiment data consid-
ers only the rotation velocity within a bout, whereas the average turning speed from the trackball includes pauses as well. Because of this, a rotation velocity of 20° may correspond to a somewhat higher left-right trackball speed in reality. Additionally, the arena experiments were conducted in the light and the trackball experiments in the dark. This has been observed to have an effect on the translation velocity of crickets [Bohm et al., 1991] although the effect on rotation is not known.

Ultimately the roughness of the approximation was not detrimental. The available phonotaxis data (Appendix A and Appendix C) gave a good indication of the distribution of turning speeds that are observed during phonotaxis, meaning that the gain of the auditory system could be fixed with reasonable certainty. The gain of the optomotor system could then be investigated over the range suggested.

### 6.8 Summary

A robot model was built, which improved on the previous robotic models of cricket behaviour reviewed in Section 2.8 in two main ways. Firstly, the trackball experiments of Chapter 4 and the close-up recordings described in Appendix C meant that more behavioural data was available to constrain the dynamics and scaling of the reflex pathways. Chapter 7 describes the behaviour of the robot when the auditory and optomotor subsystems are tested under summation at the motor output. The careful tuning of the individual subsystems to match the cricket data provides confidence that the behaviour of the robot reflects the way in which these systems would interact in the absence of a co-ordination mechanism.

The second main advance in the new robot was the inclusion of a flexible and general predictive model based upon a liquid state machine. This is justified with reference to the known anatomy and physiology of the mushroom bodies of the insect brain.
Chapter 7

Robotic Model Results

7.1 Introduction

This chapter details the testing of the robot model specified in the last chapter. The tests serve two purposes. Those in sections 7.3, 7.4 and 7.5 involve the auditory, optomotor or predictive component alone, and are primarily for validation of those systems. It is shown that the auditory component is selective for cricket calling songs with syllable repetition intervals around the species-specific value, and is capable of driving the robot towards a speaker emitting the calling song. The optomotor system is shown to produce partial correction for imposed motor asymmetries in closed-loop conditions. Testing of the predictive component concentrates on showing that it can use an auditory input to produce a reasonable prediction of the optomotor re-afference, using either off-line or on-line training. An attempt was also made to improve the performance of the model using an additional optomotor input. The problems with this approach are demonstrated.

Having demonstrated that the individual reflex pathways functioned as intended, the behaviour of the robot under summation at the motor output can be examined. It is demonstrated that the optomotor system interferes with the robot’s approach to the speaker. Since the frequency characteristics and turning speeds of the reflex pathways were tuned as well as possible to the cricket data, this is taken as evidence against the plausibility of the summation hypothesis.

The efference copy scheme is tested using both the non-directional and directional forms of the predictive model. Firstly it is shown that the two variants produce opposite effects on open-loop turning in tests equivalent to the robot feedback tests of Section 4.4.2. The two schemes are then tested in a closed-loop sound localisation task.
with random motor noise, in which they are compared to the phonotaxis sub-system alone, and the two reflex pathways summed at the motor output. It is not possible, on the basis of this test, to say that any one of the three schemes for incorporating an optomotor response are more useful to the robot.

### 7.2 Methods

#### 7.2.1 Arena

All testing of the robot was carried out in the same arena used for testing crickets, described in Chapter 5. The layout of the internal walls was left the same, and the end wall made of fabric was left in place. Black stripes, as described for Section 5.3 were fixed to the walls so that when the robot's camera was at the centre of the arena the pattern period was 60° along the horizon. Due to the height of the robot's ears (14 cm) the speakers were raised by placing them on cardboard boxes so that the centre of the speaker cones were 18 cm above the floor. Because the robot was not capable of visual object detection it was not necessary to conceal the speakers and they were placed within the main arena. Different configurations were used for different experiments. These are detailed in the relevant sections.

#### 7.2.2 Sound Stimulation

Experiments on the robot utilised both the standard synthetic calling song (see Section 5.2.2) as well as the following variants:

**Modified Syllable Repetition Interval:** For testing sensitivity to syllable repetition interval (SRI) in the auditory system the standard synthetic calling song was modified as described in [Thorson et al., 1982]. For a given SRI the integer number of syllable repetitions was chosen to keep the length of the chirp as close as possible to 250 ms. The duty cycle was fixed at 50% so the gaps between syllables also varied, serving to equalise the overall sound energy in any of the modified chirps. As for the standard song the carrier frequency was 4.7 kHz, chirps were repeated every 500 ms and there were 2 ms ramps at syllable starts and ends.

**Split Song:** A split song was used which retained the chirp structure of the standard song, but alternate syllables came from alternate speakers as described in Section 4.5.2.
Isolated Pulses: For testing the response of the recurrent pool of the predictive model a sound pattern was used that consisted of single 21 ms pulses of a 4.7 kHz tone, repeated every 500 ms from a single speaker.

7.2.3 Tracking & Path Properties

When the tracks of the robot needed to be recorded the arena’s overhead camera was used (Section 5.2.6). Tracking was performed using the same software described in Section 5.2.7. In the overhead images the robot appeared as a dark circle against the floor. Because the robot was much larger than a cricket the tracker usually located a patch containing a few tens of pixels. The uniform circular shape meant that the centre of mass of the identified patch was usually a good match for the location of the robot’s centre. In some parts of the arena, particularly when the robot was directly under the light, the estimate was poorer and was corrected by hand if necessary. Since the position but not the orientation of the robot was available in the resulting tracks they were amenable to analysis using the same path length and directness statistics used for crickets in the arena (see Section 5.2.8).

7.2.4 Stimulus Configurations

By default the robot’s motion is determined by a forward translational component, which is fixed, and a rotational component determined by the combined activities of the auditory, optomotor and predictive systems. This is referred to as closed-loop locomotion. Testing the system in detail required that the robot be operated in some more restricted configurations. Omitting the forward component results in a situation analogous to that experienced by crickets on the walking compensator (Section 2.3.4). The robot can still turn, for example to face the speaker, or to compensate for an imposed rotation. This is referred to as closed-loop rotation. During fully open-loop testing the output of the robot’s controller no longer has any effect on the movement of the wheels, although pre-programmed wheel movements may still be injected into the motor system as a convenient way to stimulate the optomotor system.

Figure 6.6 shows how the auditory and optomotor sub-systems of the robot’s controller contribute to the rotational component of its motion after multiplication by a user-determined gain. These scaling factors are usually set so the speed of movements match those seen in the insect experiments (Section 6.7), but either or both of them can be set to zero to exclude that system’s contribution on a given run. Only the output
of a sub-system is cut off in this way, so it is still possible to record and examine its internal functioning.

A few particular configurations were used in the tests described in this chapter. These are outlined briefly below:

**Open Loop Optomotor Stimulation**  In the cricket experiments of Chapter 4 open loop visual stimuli were produced by rotating the visual surround around the tethered animal. For the robot, it was simpler to rotate the robot within the visual surround. All sub-systems of the controller were disconnected from the motor output and a pre-programmed sequence of rotations was injected instead. The controller continued to compute output values and these were saved, giving the equivalent of the cricket’s intended turning on the trackball.

**Open Loop Auditory Stimulation**  In the simplest case, all systems were disconnected from the motor output, the speakers were positioned as required, and the output of the auditory system was recorded. If simultaneous optomotor stimulation was required then the robot had to rotate, and the sound field could not be physically fixed. In this case the robot’s auditory system was first used to make a recording in the correct sound field. This was saved to disk and read in by the controller in place of reading from the robot.

**Reversed Optomotor Re-afference**  This was required for replicating the feedback tests of Section 4.4.2. It was accomplished by reversing the sign of the rotational component of the robot’s motion at the output, as indicated in Figure 6.6.

### 7.2.5 Training the Predictive Model

Training the liquid state machine relies on the robot experiencing a sufficient range of correlated auditory and visual inputs. In order to collect a suitable volume of data the two speakers were placed in opposite corners of the arena as depicted in the example training tracks of Figure 7.10. The robot could then be guided around continuously by repeatedly broadcasting the calling song from one speaker, waiting for the robot to approach, and switching the song to the other speaker. In this way training sequences of a minute or more could be obtained without manual intervention being required. During a training run care was taken to drive the robot through several turns of alternating direction.
7.3 Auditory Sub-System

7.3.1 Pattern Selectivity

As discussed in the Chapters 3 and 6 (Sections 3.2 and 6.5.5) the two principal functions of the cricket auditory system are recognition and localisation of the calling song. Pattern recognition was tested using open loop conditions. The robot was placed at the centre of the arena, facing a single active speaker at a distance of 55 cm. Three different synthesised calling songs were presented to the robot in turn: the standard calling song with a 42 ms syllable repetition interval, and two variants in which the SRI had been either shortened to 18 ms or lengthened to 74 ms. For each calling song variant the activities of the model auditory neurons were recorded to disk for five seconds. Figures 7.1 to 7.3 show the first two and a half seconds of each trial.

Due to the way the auditory board enhances bilateral amplitude differences between the microphones (Section 6.4.1), coupled with a tendency for the crossing point of its output curves to be off-centre at higher volumes (Figure 6.5) there was a clear amplitude difference between the ears even though the robot was facing the speaker. The robot was left in the same place for all three recordings so the same bias is present in all three figures. It is noticeable that the peak amplitudes of the ear inputs get higher as SRI increases. This was unforeseen, and is probably due to the 2 ms ramp at the start and end of syllables in all sound patterns. This has a proportionally greater effect on the overall energy of a chirp in songs with a greater number of shorter syllables. In the trial with an 18 ms SRI (Figure 7.1) this effect has led to the input on the left hand side being too low to elicit any spikes in the left receptor neuron.

In all cases it is possible to discern the chirp pattern in the input signals (top traces). These signals are fed as simulated injected currents into the receptor neurons, which copy the pattern. When the SRI is 18 ms the receptor neuron responds with a single spike, or occasionally two spikes per pulse, and for longer SRIs a burst of spikes matching the pulse length is seen. The individual spikes are not discernible in the figure, but there are usually five per pulse at SRI 42 and ten per pulse at SRI 74 on the more strongly stimulated side. In the absence of inhibition from ON1 these spikes or bursts are relayed almost one-to-one to the AN1 and ON1 neurons, with only an occasional spike transmitted as an EPSP (most visible for SRI 18).

At the two longer SRIs the mutually inhibitory ON1 neurons interact to further en-

---

1Bear in mind the slightly higher input amplitude at SRI 74.
Figure 7.1: Auditory neuron responses to calling song with SRI 18 ms. Top traces show input to receptor neuron in pA. All other traces show membrane potential in mV.
Figure 7.2: Auditory neuron responses to calling song with SRI 42 ms. Top traces show input to receptor neuron in pA. All other traces show membrane potential in mV.
Figure 7.3: Auditory neuron responses to calling song with SRI 74 ms. Top traces show input to receptor neuron in pA. All other traces show membrane potential in mV.

All other traces show membrane potential in mV.
hance the binaural discrepancy. This can be seen by comparing the activity of the ON1 on the left side with the receptor neuron that excites it. Clearly the one-to-one correspondence between firing rates seen on the more strongly stimulated side is disrupted. At SRI 74 only one spike is fired in ON1 throughout the four chirps recorded, and at SRI 42 two spikes per chirp occur on average. At SRI 18, although there are no spikes to compare, the IPSPs that mediate the inhibition for longer pulse intervals can be seen on ON1. The inhibitory connections between ON1s and the opposite AN1s apply the binaural contrast enhancement to the ascending neurons, meaning that in all cases the chirp pattern is faithfully copied on the right AN1, but is disrupted in the left AN1.

The process essential for pattern recognition occurs at the AN1 to BN synapse. At SRI 18 ms each chirp produces a train of spikes or doublets with a modal interspike interval of 18 ms. The first couple of spikes from each of these trains produces a corresponding spike in BN1, and the synapse is then slightly depressed. Further AN1 spikes continue to occur at a relatively steady rate, causing EPSPs that keep the synapse depressed, but rarely arriving fast enough to overcome the depression and cause another spike in BN. Generally it takes a gap between chirps before the synapse can recover and the process can start again. At SRI 42 the same depression is caused by the relatively long burst of spikes received during a single pulse. Crucially, however, the inter-pulse interval is long enough that the synapse can recover. This allows the BN to fire a spike for all six pulses of each chirp. For longer SRIs BN still tends to respond with a single spike at the start of each pulse, but by definition the pulse rate is lower, resulting in a lower spike rate.

In summary, the circuit is selective for intermediate SRIs because depression in the AN1 to BN synapses only allows the BN neurons to spike at the start of a spike train in AN1. If AN1 could resolve the chirp pattern of all songs equally well this would act as a high-pass filter, firing more spikes for shorter SRIs (or, equivalently, higher-frequency pulses). However, at very short SRIs the pulses copied by AN1 are close together and chirps become indistinguishable from a single long pulse. When this occurs the BNs can only fire at the start of chirps, and rarely at other times due to fluctuations in the AN1 spike rate. This means that overall more spikes are fired at intermediate SRIs.
7.3.2 Sound Localisation

To show the ability of the phonotaxis system to locate a sound source the robot was tested in closed loop conditions. The auditory system alone was connected to the motor output with a gain of 3.22. This value, when multiplied by the open-loop output obtained in Figure 6.12 results in a mean turning speed of 20° per second. This was the value observed for turning to a sound source at 45° in the arena (Section 6.7). The translation speed was set to 4 cm/s, which is at the lower end of the path speeds observed in the close-up experiments of Chapter 5.

The two speakers were placed 73 cm apart close to one of the shorter sides of the arena as shown in Figure 7.4. The robot started each trial in the centre of the arena, facing away from the speakers. The standard synthetic calling song was played from one of the speakers only, and the neural simulation was started. Twenty trials were performed, with the side of the active speaker being switched between each trial.

![Figure 7.4: Tracks of the robot approaching a standard calling song in the arena using phonotaxis subsystem alone. Robot started at the centre of the arena, facing in the direction of the arrow. Small squares show the locations of the speakers. Ten trials shown for each speaker. Distances shown in metres.](image)

In all cases the robot approached the active speaker. The robot’s starting orientation meant that it had to perform a tight turn at first, followed by a period of maintaining a straight course. For one of the speakers two of the tracks are interrupted by the robot reaching the arena wall. The tendency to head towards this wall was seen in all runs with this speaker active. Because the layout of the arena was symmetrical
this was probably due to the off-centre crossing point of the ear sensitivity curves (see Figure 6.5).

7.4 Optomotor Sub-System

The EMD model of the robot's optomotor system contains a low-pass filter with a time constant which was tuned to give a reasonable approximation of the static characteristic curve from the cricket experiments (see Sections 6.5.4 and 4.3.2). The next stage was to tune its dynamic properties to those found for the cricket (Section 6.6.2). Having matched its properties to the cricket as well as possible based on open-loop data, how would the system behave in closed-loop conditions? The expectation was that it would achieve partial compensation for an imposed motor asymmetry, since for an optomotor controller with a static open-loop gain of $g$, the closed loop gain is given by $G = g/(1+g)$ [Collett, 1980]. The open loop gain can be read off the characteristic curve in Figure 6.10. For the motor asymmetries tested the closed loop gain falls in the range $0.61 \pm 0.02$.

To test for this effect a set of video recordings were made in which the robot started at the centre of the arena and moved on curved trajectories of different radii. These were produced by setting the robot's translation speed to 7 units (5.6 cm/s) and imposing a rotatory bias by incrementing the wheel speed on one side in conjunction with a decrement of equal magnitude on the other. Biases of $\pm 1$, $\pm 2$ and $\pm 3$ were used, resulting in rotation speeds of 17.2, 34.4 and 51.6°/s respectively. All components of the controller had zero gain, i.e. the trajectories were under purely open-loop control.

Following these initial trials a second set of recordings were made in which the test was repeated with the optomotor system active. The gain was set at $7 \times 10^{-4}$, based on the results of the trackball and arena experiments (see Section 6.7). This was judged to be at the upper end of the range of reasonable gains and causes a turning response of 100° per second when tested with the optimal pattern frequency in open loop. The resulting trajectories are illustrated in Figure 7.5.

In all cases the robot moved on curved trajectories. When the optomotor system was inactive a wheel bias speed of $\pm 1$ unit caused the robot to move on an arc with a radius of approximately 20 cm. Higher biases produced tighter circles. Activating the optomotor system compensated for some, but not all, of the bias as expected; the robot still moved on curved paths, but with a higher radius for a given bias setting.

The internal functioning of the optomotor controller when compensating for a bias
Figure 7.5: Tracks of robot in arena with motor bias and no optomotor system (solid blue lines) and with optomotor system active (speckled red lines).

is shown in Figure 7.6. Here the robot attempts to compensate for a constant rotatory bias of ±7 units whilst turning on the spot at the centre of the arena. In the top trace it can be seen that the optomotor system initially increases its output rapidly. This in turn reduces the optical flow received until a balance is reached. Note that there is not a completely steady equilibrium. This is due in part to the rotation speed of the robot being set in discrete steps, as emphasised in the middle trace. At equilibrium the mean value of the optomotor output is 4.25, compensating for 61% of the bias.

Figure 7.6: Functioning of the optomotor system with imposed bias of 7. Traces show: Top — output at summation stage. Middle — Motor output. Bottom — Resulting motion.
7.5 Predictive Sub-System

7.5.1 Internal Behaviour of Recurrent Pool

Liquid state machines are able to compute temporal functions because of the dynamics of the pool of recurrently connected neurons. Discrete inputs generate a stereotypic sequence of activity across the pool which fades over time until the quiescent state is reached. In this way a single input spike can be associated with a temporally extended output. When the input consists instead of a sequence of discrete inputs, or a continuous input function, then the pattern of activity in the recurrent pool at a given instant depends in a non-linear way on the input history.

In order to demonstrate the stereotypic response to discrete inputs the robot was placed at the centre of the arena, with the speakers at 90° to its midline. A sound pattern was presented which consisted of 21 ms pulses of 4.7 kHz sound at 500 ms intervals from the right-hand side only. With this pulse interval the facilitatory and depressive effects in the dynamic synapses of the recurrent pool would largely die away in each gap (see Section 6.5.3 for synapse model description and Appendix E for time constants). A 30-second recording of the sound pattern was made using the stationary robot’s ear circuit and saved to disk. The recording then substituted for the auditory input during the subsequent trial, in which the auditory system was connected to the motor output to allow closed loop rotation. Input to the recurrent pool of the predictive model came from the spike trains of the auditory brain neurons only, with a time delay of 70 ms. Neither the optomotor system nor the predictive model were connected to the motors, but their activity was logged.

Figure 7.7 shows the membrane potential of the right BN, the activity of the recurrent pool, and the optomotor system output during presentation of two adjacent pulses of the sound pattern. Each of the pulses resulted in a single spike in the right BN and no spikes in the left BN. Clear bursts of activity in the recurrent pool follow the BNR spikes after 70 ms. The bursts of activity last around 50 ms and occur during, or slightly before, the start of the visual response. The gross pattern of activity during the two bursts looks very similar. On closer examination it was found that the only difference between the two bursts is some jitter of the spikes fired by cell number 10 (not shown).

When the predictive sub-system is active the readout’s predictive signal must be based on the spatial pattern of activity across all the neurons of the recurrent pool. However, Figure 7.7 shows that for the parameters used in the current LSM the duration
of the burst of activity triggered by a single spike is short compared to the reafferent visual input. With regard to this problem it should be observed that during phonotaxis to the standard, six-pulse, calling song the activity of the auditory neurons is around six times higher, as illustrated in Figure 7.8. Furthermore, the readout neuron does not operate directly on the spike trains of the recurrent pool but on a filtered version. This filtering simulates the effect of integration in the readout neuron, and results in some extra blurring out of the spike trains. Figure 7.9 shows the behaviour of a random selection of the LSM filters during the same three seconds of phonotaxis shown in Figure 7.8, together with the mean activity of all the filters taken together.

### 7.5.2 Off-line Fitting of Readout Weights

The most straightforward way of testing the ability of the liquid state machine to act as a forward model was to gather a training set, and fit the readout weights using off-line linear regression. Training sets consisted of liquid states paired with corresponding optomotor feedback values, collected using the procedure described in Section 7.2.5. The ability of the model to generalise was then tested using a separate data set, collected using the same recurrent pool. This was quantified by calculating the normalised mean square error. Off-line fitting was tried for four variants of the neural model setup:
Figure 7.8: Activity of the auditory brain neurons, auditory integrator, liquid neurons and optomotor system during three seconds of phonotaxis in the arena.
Figure 7.9: Activity of five randomly chosen LSM output filters and the mean activity of all filters during the same ten-second period shown in figure 7.8. Filters correspond to neurons 63, 126, 24, 11, and 91.
Auditory input only. Inputs to the recurrent pool came only from the auditory brain neurons. The delay lines were set to 70 ms as in the previous section.

Auditory & optomotor input. The optomotor signal was also fed into the recurrent pool using two spike trains as described in Section 6.5.6. It was hoped that this would result in more continuous activity in the recurrent pool, and hence better approximation. However, it was also suspected that this arrangement might have an undesirable side-effect described below, so a control experiment that introduced a motor asymmetry was introduced.

Auditory & optomotor input + motor asymmetry. It was suspected that the inclusion of optomotor input might produce a good fit to the reafference at a given instant by simply extrapolating from the immediately preceding optomotor signal. This would mean that constant or smoothly varying ex-afferent optical flow would be cancelled out as well as re-afferent input. A constant motor asymmetry was introduced to test for this undesired effect, which turned out to be present.

Auditory input with extended delay line. Since the previous experiment showed it was not practical to include optomotor input into the recurrent pool, the auditory delay line length was adjusted manually to account for the majority of the lag in the optomotor response. This was successful in increasing the accuracy of the predictive model's approximation.

For all variations sixty seconds of training data were collected by driving the robot on a meandering path between alternating speakers. The robot was then replaced at its starting position and a further sixty seconds' worth of behaviour was captured to act as a test set. Example training and testing runs are shown in Figure 7.10.

The linear regression was performed in GNU Octave using the mldivide function. The resulting weight vector was then applied to the liquid states from the test data set to check how well the readout would generalise. It should be noted that when solving the linear regression the number of training states always exceeded the number of readout weights, i.e. the number of unknowns. If the system was truly overdetermined then mldivide would return the linear least squares solution. However it was still possible that certain states would recur often enough during a training sequence that the training data matrix was rank deficient. In this case there would be no unique solution. The
Figure 7.10: Training and testing runs used for fitting a linear readout off-line. Left — Experiment 1, auditory input only. Right — Experiment 2, auditory and opto input. Small squares show the locations of the speakers. Distances shown in metres.

behaviour of mldivide in this situation is to provide the minimum norm least squares solution, although for practical purposes any particular solution would be acceptable.

7.5.2.1 Auditory Efference Copies Only

When the recurrent pool received only auditory input the normalised mean squared error (nMSE) was 0.775 on the test data. Figure 7.11 shows the signals of the auditory integrator, optomotor system and prediction for a typical five seconds of the test run. The auditory integrator shows a series of modulations, each corresponding to a chirp of calling song. Deflections of the optomotor signal from its baseline of zero (dashed line, lower trace) occur at the same frequency due to the sound-mediated rotations of the robot. The output of the predictive subsystem (solid line, lower trace) is superimposed.

At a gross scale the predictive signal is a reasonable match for the optomotor reafference; it contains one modulation per chirp of the calling song, and these clearly have the same direction as the optomotor signal. At the scale of a single chirp, however, the predictive signal exhibits a consistent error. The start of each deflection of the optomotor signal is matched reasonably well, but the magnitude of the prediction is too low towards the end of the response. This probably relates to the duration of the sound-triggered burst of activity in the recurrent pool. Despite the extended patterns of activity that occur in response to six-pulse chirps the response of the liquid filters has decayed away by the end of each optomotor response.

It should be noted that the response of the readout is not simply a function of the
7.5.2.2 The Case for Including Optomotor Input

The inclusion of an optomotor signal, effectively an optomotor efference copy, as an input to the recurrent pool seemed reasonable given the assumption that the recurrent circuits of the mushroom bodies are a site of multisensory integration from which the extrinsic neurons must be trained to extract information appropriate to their purpose. It was assumed, in the absence of detailed knowledge, that there is no special segregation of modalities in the recurrent circuitry of the mushroom bodies. Furthermore it made sense because the motion, and hence optomotor reaference, of a cricket at a given
instant depends on the state of both the auditory and optomotor systems at previous times.

### 7.5.2.3 Auditory and Visual Efference Copies

When the input to the recurrent pool included visual as well as auditory efference copies the error in the prediction was much smaller. The nMSE on the test set was 0.191. Figure 7.12 shows a five-second period of the test run for comparison to Figure 7.11. The bursts of LSM spikes associated with chirps are longer than before, due to re-afferent visual input that overlaps with the auditory input. The improvement in the match of the prediction is visible in the lower trace.

![Figure 7.12: Prediction of LSM using auditory and optomotor input.](image)

Figure 7.12: Prediction of LSM using auditory and optomotor input. Top – bottom: Auditory integrator; LSM Spikes; Mean LSM filter activity; Optomotor reaference (dashed line) and prediction (solid line) superimposed.

### 7.5.2.4 Auditory and Visual Efference Copies + Constant Bias

As stated at the start of Section 7.5.2, above, a concern was that the a readout fed with optomotor as well as auditory efference copies would learn to ignore ex-afferent as well as re-afferent optomotor input. This section describes a control experiment to investigate this possibility.
A new pair of training and testing runs was performed using a different random connectivity of the recurrent pool. Both auditory and visual efference copies were included as inputs. Following the first test run the robot was returned to the starting point, and a second test run was performed in which a constant bias (wheel speeds ±1 causing rotation to the right) was added to the robot's motor output. Ten-second samples of the behaviour resulting from the two test runs are shown in Figure 7.13.

The left-hand side of the figure shows that when the robot experienced only re-afferent input the trained model behaved comparably to that in Figure 7.12. As expected the predictive signal was a good match for the reaference (nMSE = 0.148).

On the second test run the speaker was initially to the robot's left, and the bias was opposing the robot's tendency to steer towards it. The result was that the robot turned slightly to the right on average. Under these conditions the magnitude of the optical flow signal remained small, and the optomotor neurons that fed into the recurrent pool were largely silent (first six seconds of trace). This failure of the neurons to encode any information about the optomotor signal for low values meant that the prediction was necessarily based on the auditory inputs only, and the motor bias had no effect on it.

At around six seconds the gradual turning of the robot had resulted in the speaker being located to its rear. In this arrangement the auditory system was no longer opposing the bias, causing the robot to rotate faster to the right. This turning accelerated as the auditory system started to pick up the calling song again, this time on the right-hand side. From seven seconds until the end of the trace the robot was rotating quickly to the right, driven by the combination of the bias and the auditory system. The optomotor signal was correspondingly higher in magnitude and this was reflected in the spiking activity of the left opto neuron.

During this period the predictive signal matches the re-afferent optomotor signal closely. This is not what is required of the predictive model under the circumstances. The prediction ought to be lower, allowing the robot to recognise the discrepancy and respond to the bias. The observed behaviour indicates that the anticipated problem has occurred, and that under at least under some circumstances the readout will reflect ex-afferent optical flow in its output. On this basis it was decided to use predictions based on auditory input only for the remainder of experiments.
Figure 7.13: Two ten-second periods illustrating the behaviour of fitted forward model based on both auditory and optomotor efference copies.

Top–bottom: Auditory Integrator; left optomotor neuron (mV); right optomotor neuron (mV); spikes in recurrent pool; reafferent optomotor signal (dashed line) with prediction (solid line).
### 7.5.2.5 Setting the Delay Line Length

One way of improving the fit of an auditory-based prediction is to alter the length of the delay lines so that the burst of spiking in the recurrent pool occurs closer in time to the peak of the associated optomotor feedback. Using a delay line of 170 ms rather than 70 ms resulted in a nMSE of 0.362 during the test phase. This falls in between the previously obtained values for a prediction based on auditory input only with a delay line of 70 ms (0.775) and the flawed method with optomotor input into the recurrent pool (0.191).

### Summary

The goodness of fit obtained using the three distinct variants of the predictive model is summarised in table 7.1. Although introducing an optomotor input into the recurrent pool resulted in a better prediction the optomotor reafference, that version of the model did not allow the robot to respond correctly to ex-afferent input. The accuracy of a prediction based on auditory input only could be improved by hand-tuning the length of the auditory delay lines. This was adopted for the following experiments as the best practical alternative.

<table>
<thead>
<tr>
<th>Model Variant</th>
<th>nMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auditory + 70 ms delay</td>
<td>0.775</td>
</tr>
<tr>
<td>Auditory + 70 ms delay + opto</td>
<td>0.191</td>
</tr>
<tr>
<td>Auditory + 170 ms delay</td>
<td>0.362</td>
</tr>
</tbody>
</table>

Table 7.1: Normalised mean square error values obtained by off-line training of three predictive model variants

### 7.5.3 On-line Fitting of Readout Weights

A hypothesis put forward to explain the results of the feedback tests in the cricket requires adaptation of the predictive component over time (Section 4.5.4.3). As described in Section 7.2.5, the Widrow-Hoff delta rule can be used to train linear elements on-line. This section compares the use of the delta rule with the off-line training of Section 7.5.2. At the start of an experiment the input weights of the readout were set randomly, with a uniform distribution between -1 and +1. The delta rule was applied
at each time-step throughout a series of ten one-minute training runs conducted as described in Section 7.2.5. Figure 7.14 shows how the readout weights behaved over the course of the ten minutes.

Figure 7.14: Evolution of readout weights using Widrow-Hoff Delta Rule over ten minutes, learning rate $= 1 \times 10^{-4}$. Lower trace shows weights obtained by off-line linear regression (note different scale on the ordinate).

Inspection reveals that the relative signs and strengths of the input weights change little over time. The largest positive and negative weights at the end of the first minute remain the largest at the end of the tenth. The main change is in the absolute values of the weights. The Euclidean length of the weight vector changes from 1182 after 1 minute to 2798 after 3 minutes and 4340 after 10 minutes.
A set of readout weights for the first one-minute training period were obtained using off-line regression as before and resulted in a nMSE of 0.295. A separate regression performed for the final training period resulted in a nMSE of 0.290. By comparison, the set of weights arrived at by on-line training using the delta rule produced a nMSE of 0.578 on the first period and 0.495 on the tenth period. As would be expected, the on-line training procedure performs worse than the least squares fit. Figure 7.15 shows the values of the predictive output against the observed re-afference for the first sixty-second training period using both techniques.

![Figure 7.15: Observed re-afference vs. prediction using weights trained on-line (left) and off-line (right).](image)

The weights found by either method tend to underestimate the magnitude of the feedback at higher values. This may be due to the relative rarity of high-magnitude re-afference values in the training data, or it may reflect an inherent limitation of the non-linear mapping being performed by the recurrent pool. One interpretation of the scatter graphs is that the linear units can distinguish the direction of the input well, but the magnitude poorly. In practical terms it means that if the trained units were employed in an efference copy mechanism they would tend to cancel out only a portion of the self-generated optomotor signal when the absolute value rises above 2000.

### 7.5.4 A Non-Directional Predictive Signal

Following the trackball experiments it was proposed that crickets might maintain a prediction that takes into account the magnitude but not the direction of the re-afferent visual input (see Section 4.5.4). In order to produce a non-directional signal of the required type the readout was trained using the absolute value of the optomotor signal.
Figure 7.16 shows the fit obtained by off-line regression on the first one-minute training data set used in the previous section.

Figure 7.16: Fit of non-directional prediction to absolute value of observed reafference. Optimal readout weights calculated using off-line regression.

The scatter graph shows that the non-directional prediction has the same tendency to under-estimate the re-afference at higher optical flow values. It also acts as a reminder that the linear unit can sometimes produce a negative output. In the context of the non-directional prediction mechanism negative values are meaningless, so a (half wave) rectifying function was applied to the output to overcome this problem.

### 7.6 Sound Localisation under Summation at the Motor Output

Having tested the robot's sub-systems independently the next step was to connect the auditory and optomotor systems in parallel to examine the behaviour under summation at the motor output (Figure 2.9). At first the two systems were connected to the motor output with gains of 3.22 for the auditory system and $6 \times 10^{-4}$ for the optomotor system. The robot was examined in the same task described in Section 7.3.2. The robot started at the centre of the arena, facing in the direction shown by the arrow in Figure 7.17 (left), and the standard calling song was broadcast from one of the two speakers indicated by the small squares.

With both sub-systems connected with the given gains the initial turning of the robot was slowed by the optomotor system, and the robot's path was always interrupted by the arena wall (Figure 7.17, left). The robot still turned towards the correct side in
7.6. Sound Localisation under Summation at the Motor Output

Figure 7.17: Tracks of the robot turning towards a standard calling song in the arena, phonotaxis and optomotor systems summed at the motor output. Left — opto gain $6 \times 10^{-4}$. Right — opto gain $3 \times 10^{-4}$. Robot started at the centre of the arena, facing in the direction of the arrow in the left diagram. Small squares show the locations of the speakers. Five trials shown for each speaker at each gain. Distances shown in metres.

all cases, and tended to reach a position and orientation that would permit it to reach the speaker if the observed arc was continued. In other words, given enough space the robot would probably still have reached the speaker. However, compared to the tracks of crickets observed in the same arena, the robot’s turning was clearly restricted to curves of an unrealistic radius.

The gain of $6 \times 10^{-4}$ initially chosen for the optomotor system was at the upper end of the plausible range determined from the trackball experiments, corresponding to a maximum turning speed of 86° per second. In the next set of trials the gain was reduced to half its initial value, bringing the maximum turning velocity into line with that observed by [Böhm et al., 1991]. Figure 7.17 (right) shows the effect on the robot’s tracks. Turning is slightly tighter, but the qualitative effect is the same.

Whilst the paths of the robot are not distinguishable from smooth curves when plotted in the figures, examination of the internal functioning of the simulator reveals finer details. Figures 7.18 and 7.19 show representative examples of the auditory and optomotor system outputs during one track from the left and the right plots respectively of Figure 7.17. From these plots it can be seen that the chirp pattern of the calling song is represented in the motor output of the robot, and that the optomotor system is capable of resolving and responding to these individual turns to sound. Frequently the optomotor reaction is large enough to cause over-compensation, where the sign of the steering signal transiently reverses in sign. This is seen throughout the motor output
traces in Figure 7.19 and Figure 7.18.

![Graph](image)

**Figure 7.18:** Outputs of the auditory and optomotor systems during turning towards a sound source with optomotor gain of $6 \times 10^{-4}$ (upper two traces) and resulting motor command after summation and rounding (lower trace)

### 7.7 Comparison to Feedback Tests on Trackball

The surprising result of the feedback tests on crickets (Section 4.4.2) was the difference in the cricket's turning in the two open loop conditions, dependent on the sign of the previous optomotor feedback (Figure 4.16). This led to the hypothesis that the crickets might employ a *non-directional* predictive model, since the more usual directional form of an adaptive efference copy was expected to result in the opposite curve ordering (Section 4.5.4). Experiments were conducted with the robot to check that the implemented predictive models really did behave as proposed in tests equivalent to the insect experiments on the trackball, and that the effect on the curves could not be explained by summation at the motor output.

To achieve comparable stimulus conditions the robot had to receive open-loop sound stimulation at the same time as either closed loop visual feedback (positive or negative) or zero visual feedback. For the sound stimulation a twenty-second recording was made using the auditory system of the stationary robot. As for the cricket experiments the standard calling song was played for fifteen seconds from one side only (the right), followed by 5 seconds of silence. During a trial the recorded pattern was read in repeatedly in place of microphone readings.
Figure 7.19: Outputs of the auditory and optomotor systems during turning towards a sound source with optomotor gain of $3 \times 10^{-4}$ (upper two traces) and resulting motor command after summation and rounding (lower trace)

The robot was made to respond to the calling song for twelve cycles of the twenty-second sound recording. Each time the sound loop was started the optomotor feedback situation was changed in the following sequence: positive feedback – open loop – negative feedback – open loop.

### 7.7.1 Summation at the Motor Output

The first experiment was conducted as a control, to check that the behaviour of the crickets described in Section 4.4.2 could not be produced by the simpler model without a predictive component. Figure 7.20 shows the auditory, optomotor and motor signals over one iteration of the four feedback situations.

During negative feedback the behaviour of the motor signal is the same as that demonstrated during closed-loop phonotaxis (compare to Figures 7.19 and 7.18). The optomotor system opposes but does not completely negate turning towards the sound. The motor output sometimes takes negative values, when the optomotor system causes the robot to turn back slightly.

Positive feedback causes a stable elevation in the magnitude of the optomotor signal. When calling song is present the optomotor feedback increases the size of the turn to each chirp. When the calling song stops the level of turning decreases, but the feedback sustains it at a low level until the visual situation switches to open loop. When the switch occurs the optomotor output dies away, so that by the time the first auditory
Figure 7.20: Auditory, optomotor and motor output during one iteration of the feedback test sequence. Auditory and Optomotor systems summed at the motor output. Vertical dashed lines indicate change of feedback state.
7.7. Comparison to Feedback Tests on Trackball

7.7.2 Directional Predictive Model

The test described above was repeated with the directional predictive model active. First the model had to be trained using the procedure of Section 7.2.5, then the robot was placed at the centre of the arena and the feedback test was started. During the test the learning rate of the predictive model was set to $1 \times 10^{-4}$.

When the resulting motor output was averaged across the chirp period for each feedback situation (as in Figure 4.16) the curves of Figure 7.21 were produced. These show that the robot responded differently in the two open-loop situations, but the order of the curves does not match that of the crickets as shown in Figure 4.16.

![Figure 7.21: Motor output averaged across the chirp period in four feedback situations for adaptive directional forward model](image)

Figure 7.22 shows how the ordering of the curves arises. The eighty second period shown in the diagram is the third and last iteration of the test. The auditory signal is identical in each feedback situation, since it is being read in from the same file.
each time. The optomotor signal is negative during negative feedback, positive during positive feedback and zero during open loop phases; it is the prediction alone that accounts for the difference between the response in the two open-loop blocks.

![Graph showing auditory and optomotor signals, prediction, and motor output during the third iteration of the feedback test sequence.](image)

**Figure 7.22:** Auditory and optomotor signals, prediction and motor output during the third iteration of the feedback test sequence. Directional predictive model with learning rate $1 \times 10^{-4}$. Vertical dashed lines indicate change of feedback state.

The third trace of the figure shows how the predictive signal produced in response to sound fluctuates throughout the experiment. During negative feedback, the prediction is driven from positive towards negative values. Therefore, during the subsequent open-loop block, the range of the predictive signal is approximately $-0.65 - 0.85$. The absence of visual feedback at this stage means that the prediction values will be driven
closer to zero. The rate of this change is slow, however, because the training signal is proportional to the error and the absolute value of the signal are already small. When positive feedback begins the optomotor feedback from a chirp is always positive, and this causes the prediction to be driven towards positive values. Towards the end of the positive feedback block the range of the predictive signal is approximately -0.12–2.34. Since the predictive signal is subtracted from the optomotor signal when the two are combined the resulting response to a chirp is smaller during the post-positive open-loop phase than during the post-negative open-loop phase.

### 7.7.3 Non-Directional Predictive Model

The final feedback test used the non-directional form of the predictive model. As for the directional model an eighty-second training run was performed, this time using the absolute value of the optomotor feedback as the target of the linear regression. Over the course of the subsequent tests the prediction was allowed to adapt with learning rate $5 \times 10^{-4}$. Averaging the recorded motor signal over the chirp period for each feedback situation gives the curves of Figure 7.23.

![Figure 7.23: Motor output averaged across the chirp period in four feedback situations for adaptive non-directional forward model](image)

Using the non-directional model the order of the curves for the two open-loop blocks matches the order seen in the cricket experiments. Figure 7.24 shows how the ordering of the curves arises from the adaptation of the predictive signal over time.

In the non-directional model the predictive signal (third trace of Figure 7.24) should,
Figure 7.24: Auditory and optomotor signals, rectified prediction, discrepancy and motor output during the third iteration of the feedback test sequence. Non-directional predictive model with learning rate $5 \times 10^{-4}$. Vertical dashed lines indicate change of feedback state.
under ordinary circumstances, match the magnitude of the optomotor response. The magnitude of the actual response is subtracted from the prediction to give the discrepancy (fourth trace). It is this discrepancy signal that is injected into the motor pathway to try and compensate for the observed mismatch. In the feedback test the turning tendency of the robot is always towards the sound which is to the right (positive in the graph), so the discrepancy trace is always added to the auditory and optomotor traces in the example.

The difference between the two open-loop conditions arises because the magnitude of the optomotor input during positive feedback is always larger than that received during negative feedback. This drives the the predictive signal to a higher value at the start of the post-positive block compared to the post-negative block. During these open-loop blocks the absence of optical flow means that the robot tries to turn faster to compensate. The magnitude of this turning is larger in the post-positive case, corresponding to the larger predictive signal.

7.8 Behaviour Under Random Motor Disturbances

It is presumed that the purpose of an optomotor response is to compensate for external disturbances to the movement of the robot. Therefore the final experiment examined the behaviour of the robot under conditions of random motor disturbance. A rotatory bias, taken from a uniform distribution on [-2,2], was superimposed on the robot’s motor output. Throughout each trial the bias value was re-sampled at random intervals with a mean interval of one second.

The ability of the robot to track the source of a calling song in the presence of this disturbance was tested using:

- The phonotaxis controller alone.
- The phonotaxis and optomotor systems summed at the motor output with gains of 3.22 and $3 \times 10^{-4}$ respectively.
- The phonotaxis and optomotor systems as described above, plus the directional predictive model.
- The phonotaxis and optomotor systems as described above, plus the non-directional predictive model.
The speaker positions and starting orientation of the robot were changed to that shown in Figure 7.25. With this layout the optomotor system would be unlikely to prevent the robot from reaching the speaker, allowing the tracks produced under the different conditions to be compared. Ten test runs were carried out for each configuration of the controller. If a predictive model was used then the robot underwent a 60-second training run before the trial. Linear regression was used to set the readout weights, which were then fixed for the duration of the experiment.

It was expected that in all cases the introduction of the optomotor system would help the robot correct for external optical flow and move to the speaker with less meandering. Under summation at the motor output the optomotor response was expected to slow the robot’s turns to sound and result in more curved paths, whereas the predictive models were expected to correct for this, and produce the most straight, direct paths.

### 7.8.1 Results

Figure 7.25 (top-left) shows the behaviour when the phonotaxis system alone controls the robot, and Figure 7.25 (top-right) shows the effect of adding the optomotor system. In both experiments the robot approached the correct speaker in all trials, although the spread of the finishing positions was much greater than that seen in the absence of random turns (Figure 7.4). Sometimes the robot stopped when it reached the casing of the speaker, but sometimes it missed the speaker and instead reached a nearby part of the arena wall. In one case in the left-hand figure the trial ended before any obstacle was reached. Qualitatively it appears that the introduction of the optomotor system caused the tracks of the robot to become straighter. This is particularly apparent for those tracks approaching the speaker on the right-hand side of the diagram.

Figure 7.25 (bottom left) shows the behaviour of the robot using a directional prediction and the bottom-right panel shows the effect of a non-directional prediction. The tracks with the directional prediction show a surprising asymmetry, with those approaching the right-hand speaker looking particularly regular, but those approaching the left-hand speaker indistinguishable from those obtained with the auditory system alone (Figure 7.25). None of the tracks obtained using the non-directional prediction look particularly different in character to those obtained in the phonotaxis-only experiment. The one track that crosses the centre-line of the arena was recorded when the right-hand speaker was active.
Figure 7.25: Tracks of the robot turning towards a standard calling song in the arena under conditions of random motor disturbance. Top-Left — Phonotaxis system only. Top-Right — Phono and opto systems summed at the output. Bottom-Left — Directional prediction. Bottom-Right — Non-directional prediction. Small squares show the locations of the speakers. Five trials shown for each speaker in each experiment. Distances shown in metres.
7.8.2 Predictive Models & Summation Compared

In order to make a quantitative comparison between the tracks produced by the various controllers under random disturbance the directness statistic D (see Section 5.2.8) was calculated for each track. The values obtained are plotted in Figure 7.26. From the graph it is apparent that the sets of D values overlap to a large extent. A one-way ANOVA was conducted to test whether the variation across the different controller types was significant. The ANOVA table in Figure 7.2 shows the that probability of the observed variation in the D statistic, assuming the null hypothesis that the controller makes no difference to the path directness, is 0.49. The null hypothesis cannot be rejected.

![Figure 7.26: Obtained values of the D statistic for sound localisation runs under random motor disturbance. Left – right: Auditory only; Auditory & optomotor systems summed at the motor output; Directional predictive model; Non-directional predictive model. Ten tracks measured for each controller variant.](image)

Surprisingly, neither of the two predictive controllers (in non-adaptive form) were successful at producing more direct paths to the speaker under conditions of random motor disturbance. The paths produced by the non-directional controller just appear to be very variable. On examination of the tracks produced by the directional controller there is an evident qualitative difference between the tracks on the two sides of the arena. The right-hand tracks form a cluster with D values around 0.4, and all the left-hand tracks have lower D-values.

This discrepancy between the two sides of the arena also appears to be present in the tracks for summation at the motor output, but it is not obvious in the tracks for the
non-directional controller. It is hypothesised that the presence of a cloth wall at the right-hand end of the arena led to fewer echoes there, making the sound localisation task easier in the vicinity of the right speaker. It does not, however explain why the efference copy controller should have performed worse than the phonotaxis controller alone at the left-hand end.

7.9 Discussion

7.9.1 Fit of the Robot Model to the Cricket Data

Auditory System

The assumptions made in the design of the auditory board have already been discussed in Section 6.4.1. The auditory neural network is very similar to the one from [Reeve and Webb, 2003], except that only one stage of brain neurons was used. As demonstrated in Section 7.3 the BN neurons alone still show band-pass properties when tested with sound patterns like those of [Thorson et al., 1982]. Omission of the BNC2 neurons from [Reeve and Webb, 2003] means that turning to songs with the longest SRI values is not fully suppressed in the current model, but since the multimodal experiments only use standard calling song this difference is not critical to their evaluation.

A further difference between the current model and its predecessor is the use of a different spiking neuron model, namely Izhikevich's quadratic model [Izhikevich, 2007a]. Successful use of the alternative model in pattern recognition and sound localisation indicates that these properties arise from the network structure and are not a consequence of the integrate-and-fire model used previously.

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>Prob&gt;F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Columns</td>
<td>0.0321</td>
<td>3</td>
<td>0.0107</td>
<td>0.82</td>
<td>0.4929</td>
</tr>
<tr>
<td>Error</td>
<td>0.47142</td>
<td>36</td>
<td>0.01309</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0.50352</td>
<td>39</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7.2: ANOVA table for one way uncorrelated ANOVA on D values from the four controller variants (Auditory only, Auditory + Opto, Directional Prediction, Non-directional Prediction).
When the auditory system alone was operated in closed-loop conditions the robot tended to produce smooth paths towards the speaker (Figure 7.4). For lateral positions of the speaker the path was a curve, and for frontal positions the robot headed roughly straight. This contrasts with the paths of crickets in the arena experiments, which tended to meander towards the speaker whether in the light (e.g. Figure 5.3) or the dark (Figure 5.8). Given the rapid steering capabilities of crickets on the trackball (Section 2.4.3) it is a little surprising that crickets do not go equally straight. Possibly the robot ears are more accurate at resolving the angle of azimuth of the speaker. This could be due to the fact that the cricket’s ears are on its front legs and move about as it walks.

In Figure 7.4 the robot was deliberately tested by starting it in a position where it faced away from the speaker. It is noteworthy that there was rarely a trial in the arena where a cricket started off heading away from the speaker. Usually by the time a cricket left the piece of egg box on which it started it was heading into the correct half of the arena. Figure 5.5 (left) shows two examples of crickets which did not do so, one of which initially describes a smooth curving path. It would be interesting to see if more paths of this nature could be obtained using a speaker-switching paradigm like that used for training the robot’s prediction. In this way the sound could be switched to a position lateral to the cricket. If this produced smoothly curving paths it would support the idea that the cricket was making continuous corrections, as predicted by the rapid steering model, rather than turns of a fixed size [Oldfield, 1980] or alternate turns and straight walking bouts [Murphey and Zaretsky, 1972].

Optomotor System

Tuning the delay filter of the EMDs in the optomotor system to match the cricket’s characteristic curve (Figure 6.10) also resulted in a reasonable fit of the frequency response (Figure 6.14) and the step response (Figure 6.16). As stated in Section 6.6.2, the phase/frequency match is largely fortuitous, since there is a lag of approximately 150 ms inherent to the robot’s video system (see step response, Figure 6.16), which is twice as long as the lag estimated for the cricket data (Figure 4.13).

Despite the extra fixed lag in the robot’s optomotor system, the frequency characteristics of the EMDs meant that for frequencies below 4 Hz the phase lag of the optomotor response was actually around 0.5° less than that of the cricket. At the chirp rate of the standard calling song (2 Hz) this amounts to 40 ms. This means that the tendency of the optomotor response to cause over-compensation during chirp gaps, as
seen in Figures 7.18 and 7.19, is probably somewhat reduced in the robot compared to what would be expected in a cricket, assuming summation at the motor output.

Feedback Tests

Further indication of some discrepancy between the dynamics of the two reflex pathways and those of real crickets comes from the feedback tests in Section 7.7. When the responses of the robot over the chirp period (Figure 7.21 and Figure 7.23) are compared to the equivalent curves from the cricket experiments (Figure 4.16) there are some obvious differences. In the robot experiments the syllable pattern of the calling song is evident, which is not the case for the crickets. In addition the influence of the optomotor feedback is only evident during the last 300 ms of the chirp period. Between zero and two hundred milliseconds the curves for the different conditions look almost identical, apart from the top curve in Figure 7.21 which is different due to the self-sustaining positive feedback. The cricket data, meanwhile shows approximately equal shifts at all phases of the chirp period.

These differences suggest that the dynamics of the robot’s reflex pathways may not have fully captured the characteristics of their biological targets, and that the time constants of filtering in both systems ought to have been longer. Slower dynamics in the auditory system would smooth out the effect of individual syllables. Slower dynamics in the optomotor pathway could act to blur together the responses to individual chirps, creating a relatively steady shift as seen in the cricket.

There are other possible explanations for the differences: The cricket’s responses to chirps on the trackball are quite variable (see for example [Hedwig and Poulet, 2004], Figure 1B) and do not obviously copy the syllable pattern. This may be due to noise in the system, or interaction with the cricket’s stepping cycle, which causes modulations at around 6 Hz ([Hedwig and Poulet, 2004], Figure 1C). When many chirps are averaged together these uncorrelated components will be lost. The blurring of the optomotor response to create a constant shift might have been due in part to slow dynamics of the motor driving the patterned cylinder.

If the dynamics of the robot’s systems were made slower, then accuracy with respect to other measurements (auditory rapid steering amplitude or the half-time of the optomotor step response) would become worse. Since the aim of the robot feedback tests was to demonstrate production of the feedback-dependent adaptation observed in the cricket (which was achieved) then the current compromise is acceptable.
7.9.2 Performance of Predictive Model

Suggested by the circuitry of the mushroom bodies, the idea of the predictive model was to produce a sustained response to discrete inputs via recurrent connections. In the absence of information that could allow an accurate model of the connectivity of the mushroom bodies to be constructed the generic "cortical" wiring of [Maass and Natschläger, 2002] was used.

Examination of the predictive model output when disconnected from the motor loop demonstrated that the liquid state machine was capable of producing a reasonable prediction of the time-varying magnitude of the optomotor reafference during the burst of recurrent activity that resulted from a chirp (Figure 7.11). However, the recurrent activity tended to die out rather abruptly, resulting in relatively long periods during which there was no activity in the recurrent pool and hence zero output. A practical solution to this problem was to tune the delay line lengths by hand so that the activity in the pool was aligned with the time when the optomotor reafference was at its greatest.

In theory the liquid state machine ought to be able to learn the delay of the optomotor reafference, however this assumes that the activity in the recurrent pool is continuous. When the activity is driven by an input that is inherently discontinuous this is not possible, unless the recurrent excitation in the pool can sustain itself for long enough to last throughout the gaps. Initially it appeared that visual input into the recurrent pool might overcome this problem. However, the inclusion of optomotor efference copies resulted in the LSM learning a trivial self-correlation, which had no predictive power.

This demonstrated a limitation of the neural architecture. If sensory predictions are generated by circuitry like an LSM, then the extrinsic/readout neurons that produce predictions of a given modality should not contact Kenyon cells providing information about the same modality. This implies that different modalities should be segregated in the Kenyon cells, and that the gross associations between different modalities should be pre-determined in the mushroom body structure. Learning would only tune the temporal relationships.

---

2Zero output in the absence of a bias weight on the readout unit. In the presence of a bias weight (not shown) the readout unit tends to predict the average value of the optomotor re-afference, which is approximately zero depending on whether the training run was biased.
7.9.3 Implications for Mechanism of Co-ordination

Summation at the Motor Output

As discussed in Section 7.9.1, the phonotaxis controller of the robot tended to produce rather straight paths to the speaker once it was facing the right way. During these types of movement the optomotor response would have little effect. The best evidence that the optomotor response would interfere with turning comes from movements when the robot is turning to face the speaker, as shown in Figure 7.17. Even when the optomotor gain is set to a moderate level of $3 \times 10^{-4}$ the robot is clearly restricted in its movements when compared to Figure 7.4. However, as also discussed in Section 7.9.1, this type of turning was not representative of the sort of turning seen from crickets in the arena. This suggests that summation at the motor output is not necessarily a bad strategy for a cricket, which is borne out by the testing of the controller under conditions of random motor disturbance. Starting at an angular deviation from the speaker of roughly $30^\circ$ the robot was still able to reach the speaker, although without a significant increase in directness.

The fact that crickets in the arena managed to align themselves with the speaker before beginning their phonotactic run suggests that this early phase of orientation ought to be interesting for future investigation. In the experiments of Chapter 5 the crickets were placed into the arena at no particular orientation. This means that they must have had to align themselves by turning on the spot while standing on or under their piece of egg box. Other arena studies have reported “scanning” behaviour, in which a cricket spends some time (5–10 seconds in [Murphey and Zaretsky, 1972]) rotating from side to side on the spot before starting to walk. If Gryllus could be made to reliably perform rotatory scanning behaviour it would be interesting to obtain some film in light versus dark conditions, since during this behaviour in the light they would receive self-generated optomotor input consisting of almost pure rotation.

An indication that summation at the motor output is not the strategy adopted by the cricket comes from the close-up recordings in the arena. There, crickets were observed to pause before commencing turning at the start of the next bout. Over the course of a walking bout average turning speeds of $20 – 40^\circ/s$ were observed, implying even higher instantaneous speeds (Figure C.7). Given the behaviour of the model optomotor controller in response to movements of around $20^\circ/s$ (Figure 7.19), what is known about the lag of the optomotor response compared to the model (Section 7.9.1), and the comparable duration of a walking bout to the turns made in Figure 7.19, a strong
overcompensatory response might be expected. These are not seen.

**Efference Copy Mechanisms**

The evidence presented above does not definitely require, or rule out, an efference copy mechanism in the cricket. The optomotor response does not provide complete compensation for visual motion, so summation at the motor output would not genuinely "immobilise" the insect. However the best available evidence suggests that it would indeed interfere with intentional turns, and the arena experiments do not show the expected overcompensation.

Two variants of efference copy scheme were tried on the robot. The directional version implemented the classic efference copy scheme of Figure 2.13, while the non-directional version was created to try and account for the results of the cricket feedback tests of Section 4.4.2. When adaptive versions of the two schemes were tested under conditions equivalent to that test the correct curve ordering was produced by the non-directional controller, while the directional controller produced the opposite, incorrect, curve ordering. This is not a reason for rejecting the directional model, however, since the correct curve ordering could be arrived at by assuming the presence of a sustained after-effect of optomotor stimulation. If this is the correct explanation, then the after-effect is not captured in the model optomotor system, as demonstrated in Figure 7.20.

When examined off-line the output signals produced by the two predictive models seemed qualitatively reasonable. The subsequent lack of a significant improvement in path directness when using the predictive models under random motor disturbance is surprising. In [Webb and Reeve, 2003] inhibition was put forward as the most likely of the schemes considered there, on the basis of its success at stabilising the robot. No such argument can be made for any of the optomotor controllers presented here.

A reason to prefer the directional optomotor controller over the non-directional variant is its neural plausibility. Based on what is known about the cricket auditory system it is expected that the turning rate of the cricket will be encoded in the activity of bilateral neurons, for example in the difference of their firing rates. A liquid state machine-like network could be adapted without difficulty to encode the expected optomotor re-afference in a bilateral pair. Combining this signal with the cricket's descending pathways could be effected by a single excitatory or inhibitory connection on each side. The non-directional controller, meanwhile, has to apply its non-directional discrepancy signal according to the current direction of turning. In the current model this is determined by relatively sparse spikes, and can change rapidly. It is difficult to
conceive of a simple circuit that could mediate this selection process on the required time-scale.
Chapter 8

Summary and Conclusions

The aim of this thesis was to investigate the co-ordination of multimodal behaviour by studying a specific example: phonotaxis and optomotor following in the cricket. These behaviours were relatively well understood individually, but there was only a very little known about their interaction. Five hypotheses were identified to guide a series of behavioural and robotic experiments.

8.1 Behavioural Clues to the Mechanism of Interaction

The time-averaged tests of Section 4.4.1 replicated the experiments of [Böhm et al., 1991] with a greater number of crickets, to test whether the direction of a calling song had any effect on the magnitude of the optomotor response. No effect was found. Comparison to the magnitude of the optomotor response in vision-only tests also revealed no significant change. On this basis it was concluded that the inhibitory scheme proposed in Section 2.6 did not apply to the cricket optomotor response, and neither was the gain of the optomotor system being modulated in any significant way in the presence of calling song from any direction.

In order to improve upon previously built models it was necessary to match the dynamics of the visual system to the cricket, rather than to Drosophila. The open-loop trackball was used to measure the static characteristic curve and frequency response of the cricket’s optomotor following, and this was used to tune the lag parameter of an array of elementary motion detectors. This also produced a reasonable approximation to the frequency response observed in the cricket.

The information generated during the characterisation of the optomotor response allowed the feasibility of the chained subsystems model (Section 2.6.3) to be assessed.
Whilst it is not possible to be categorical about whether the auditory signals in the brain feed through the optomotor system, the high frequencies present in split songs are not attenuated as much as would be predicted given that arrangement. Following the trackball experiments (and some simple computer experiments with split songs and filters) it was thus apparent that the additive or efference copy hypotheses were the most likely of the five originally proposed.

Distinguishing between summation at the motor output and efference copy using behavioural experiments is not possible unless the efference copy pathway contains an adaptive component. The arena experiments of Chapter 5 should have revealed the presence of an additive mechanism or fixed efference copy as opposed to an adaptive mechanism. Responses to individual elements of the striped visual pattern confounded the results somewhat, but there was no firm evidence to suggest that path directness was affected by attempts to change the insects' optomotor feedback. The trackball-based feedback tests of Section 4.15 also suggested the possibility of the presence of an adaptive efference copy, although the effect seen could also be attributed to slow dynamics in the visual system.

### 8.2 Insights From the Robotic Model

A robotic model was constructed to explore the behaviour of the two favoured hypotheses and try and distinguish between them. The dynamics of the two reflex systems were tuned to the data from open-loop trackball experiments.

Despite this tuning the behaviour produced by the phonotaxis system in closed-loop conditions was not cricket-like; the paths were too smooth. In previous models meandering would have been produced by turns of fixed speed and duration. These ideas are not compatible with the most recent cricket data, so it seems more likely that the meandering seen in crickets would be a product of coarse angular resolution of the auditory system in the frontal field.

When the auditory and optomotor systems were summed at the motor output the optomotor controller interfered with the robot's turns to sound, for a plausible range of gains (Section 7.6). Turns of crickets in the arena did not seem to display the same kind of overcompensation. However, the optomotor response would not prevent the robot from reaching the speaker, so the summation hypothesis could not be ruled out. The summation hypothesis alone would not account for the results of the feedback experiments of Section 4.4.2 unless the model visual system did not fully capture the
dynamics of the optomotor response as suggested above.

No previous model of auditory and optomotor interaction had incorporated a full efference copy mechanism, only directionally predictive inhibition. Inhibition having been discounted as a possible mechanism, the new model used one of two simple algorithms to combine the output of a predictive model with the output of the optomotor pathway. The predictive component itself was a liquid state machine, based on the idea that the recurrent circuitry of the mushroom bodies might act to generate cross-modal associations and sensory predictions.

The first version of the model generated a directional prediction of the optomotor re-afference and combined it using a simple summation. The second version was created to try and explain the results of Section 4.4.2. It generated a non-directional prediction of the magnitude of the optical flow. The discrepancy between the prediction and the received re-afference was then used to alter the speed of turning, without changing the direction. It is not possible to express a preference for one of the two models based upon their ability to correct for random turns in closed loop conditions. Neither one performed significantly better than the phonotaxis controller alone. This is surprising, given the success of an efference copy-like inhibition mechanism in a previous study [Russo et al., 2005]. It appeared that the robot’s ability to navigate was adversely affected by echoes in one part of the arena, causing a large variability in directness scores.

Whilst only the non-directional model could account for the feedback test results, it is not particularly plausible when considered from the point of view of constructing neural networks.

8.3 Mechanism of Combination — Summary

On the basis of the behavioural experiments there is no evidence for accepting the inhibition or gain modulation hypotheses, and good grounds for rejecting the chained sub-systems hypothesis. Further investigation of the summation and efference copy hypotheses have shown that:

- According to the best available parameterisations of model reflex pathways the optomotor response would interfere with phonotaxis under summation at the motor output.
- A directional efference copy scheme, employing on-line adaptation via feedback
error learning, would not produce the ordering of the curves shown in the feedback tests. However, if the visual system were to show sustained after-effects of optomotor stimulation then this would account for the results of the feedback tests, while the efference copy (which need not be adaptive) could account for the ability of the crickets to orientate rapidly in the arena.

- A non-directional efference copy can qualitatively reproduce some aspects of the feedback tests. However the algorithm created to fit the data is not particularly plausible from the point of view of constructing a neural network.

On the available evidence it is not possible to decide conclusively between the three options. The second possibility mentioned here — directional efference copy plus visual adaptation — is the most elegant, although as [Wehner, 1987] emphasises, evolution does not select for elegance. The non-directional mechanism implemented is rather speculative, and while it does provide a fit to the behavioural data there is little evidence to support its utility for course stabilisation. Implementation of summation at the motor output only suggests that there is nothing special about the dynamics of the two systems that would allow them to co-exist without interaction, but there are still uncertainties about whether the dynamics of the two systems were captured correctly. Bearing these caveats in mind it is tentatively proposed that the cricket employs an non-adaptive directional efference copy to predict visual re-afference.

### 8.4 The Liquid State Machine as an Analogue for the Mushroom Bodies

The results of [Schildberger, 1984a] and [Mizunami et al., 1998a] indicate the existence of sustained, motor-linked, activity in the mushroom body extrinsic neurons, which could represent sensory predictions. In the case of Schildberger's recordings the duration of the activity could only reasonably be produced by recurrent activity.

It is impossible to say, at present, whether the Kenyon cells and extrinsic feedback neurons of the mushroom body act like the recurrent pool of a liquid state machine. In order to make this claim it would be necessary to show that the particular structure of the mushroom body is capable of producing sustained sequences of activity that have the separation property [Maass and Natschlager, 2002] (see also Section 6.3). Demonstrating this in vivo would involve simultaneously recording from many axons of the
pedunculi. Tetrode recordings in the locust mushroom bodies have been used to isolate and examine individual Kenyon cell responses to odours [Broome et al., 2006]. However these responses are not very informative for current purposes, because olfactory inputs arise from the antennal lobe which is itself highly recurrent and generates sustained responses. The author is not aware of any such recordings that examine responses to non-olfactory stimuli. In any case, even a tetrode would only be able to resolve a tiny fraction of the whole Kenyon cell population at a time.

For the present, the most fruitful approach to answering this question would be to take the properties of a number of individual, hopefully representative, cells and demonstrate the separation property in a model as Maass and colleagues did for the cortex. The electrophysiological properties of Kenyon cells have already been determined in some detail in the bee [Wüstenberg et al., 2004], and the body of evidence regarding extrinsic neurons and their connectivity is growing [Tanaka et al., 2008], so more accurate models are likely to become feasible.

Constructing and testing the cricket model highlighted an expected feature of a cross-modal association area: gross associations should be pre-determined. When this was not enforced in the model the visual input from the optomotor system was associated with the visual feedback at the next time-step, creating a useless self-association (Section 7.5.2.4). It is known that the pedunculi and lobes of the mushroom bodies contain lamina that carry information from different modalities and that certain extrinsic neurons at the lobes have processes in restricted regions [Tanaka et al., 2008]. It is possible that some of these sub-divisions represent pre-determined cross-modal associations.

Overall the analogy between the mushroom bodies and a liquid state machine is a highly speculative one. However, the evidence for sustained, motor-related activity in the protocerebrum of crickets and cockroaches is sound. The presence of local recurrent circuitry to support this activity is highly likely, even if it turns out not to involve the Kenyon cells.

8.5 The Liquid State Machine as a Substrate for Sensory Predictions

The anticipation and cancellation of predictable sensory re-afference is an important and general task of nervous systems. In essence the problem is one of computing a non-
linear mapping from one continuous signal (an efference copy, or conditioned stimulus in a classical conditioning task) to another (the sensory re-affle rence, an unconditioned stimulus, a proprioceptive signal). Construction of the multimodal cricket robot has shown that a liquid state machine is a practical and biologically plausible method for accomplishing this function, although the discontinuous nature of the input means that fixed delay lines had to be introduced rather than allowing the LSM to learn the delay.

In order to learn the delay the activity in the recurrent network would have to be continuous. Increasing the size of the recurrent network is one way of increasing the memory of a recurrent pool [Maass et al., 2004], although it was not an option given the hardware used in this study; the neural simulator was already operating at the limits of the off-board computer. At present there is no rigorous theory of how to alter the parameters of a spiking neural model with dynamic synapses to obtain different dynamics, although a theoretical analysis of similar randomly connected circuits of simple threshold gates is possible [Bertschinger and Thomas, 2004].

Liquid state machines have previously been trained to perform sensor–motor mappings in simple obstacle avoidance and steering tasks on small robots such as the Khepera [Burgensteiner, 2006] [Dutoit et al., 2007]. As far as the author is aware they had not yet been applied to a practical sensory prediction task. Joshi and Maass present the use of an LSM for controlling the trajectory of of two-joint robot arm based on position and velocity feedback as well as “efference copies” of the motor command at the previous time step [Joshi and Maass, 2005]. This was only tested in simulation, however.

### 8.6 Future Work

#### 8.6.1 Cricket Behaviour

Building and testing the robot has highlighted several aspects of the cricket’s behaviour which it was beyond the scope of this PhD project to examine in detail, but which may contribute to phonotaxis in the presence of visual cues. The influence of leg movements on the directionality of hearing is rather poorly understood. Ongoing attempts to model the legged locomotion of crickets may lead to the possibility of a descriptive model of the time-varying ear position, or even the possibility of placing microphones on the legs of a hexapod robot. This may help explain why the meandering paths of crickets are produced, even when orientation on the trackball is so precise.
High speed video recordings have revealed that crickets orientating to sound on the trackball make movements of the head and thorax, but this is not seen in the arena (G. Petrou, personal discussion). These movements are unlikely to make any difference to the time-averaged performance of a cricket in a protocol like that of Section 4.4.1, but may be more important over short time scales, for instance during rapid steering (Section 2.4.3). Incorporation of the influence of these head movements would increase the accuracy of the model.

Comparison between the robot model and the tracks of crickets in the arena highlighted the fact that the crickets rarely started their run towards the speaker until they were facing in the right direction. This orientation in place, or 'scanning' seems to be a component of the phonotaxis behaviour in *Gryllus* and other species, and is likely to be of interest because the cricket would experience rotational flow alone. In this situation the type of model optomotor response used on the robot is certain to be applicable.

### 8.6.2 Separating Exafference from Reafference in Learning in the Robot

The ability of the liquid state machine architecture to learn was, in the current model, only exploited in the robot feedback tests. There, the only visual input experienced by the robot was re-afferent. In an insect's natural environment the optomotor signal will contain ex-afferent components too. The current learning rule, which doesn't discriminate between the two, could still produce a useful prediction if the ex-afferent signals were unbiased and of short duration. However the current rule would learn a long-term bias which is undesirable.

A simple rule for determining whether input is ex-afferent or re-afferent is that re-afferent signals should be temporally correlated with the efference copy. This being the case the predictive component could utilise a learning rule that employed STDP or some other kind of fading synaptic marker (e.g. [Izhikevich, 2007b]) to respond only to re-afference. STDP has been shown to exist between Kenyon cells and mushroom body extrinsic neurons, although in the locust it seems to exist to fine tune the phase of firing in the extrinsic neuron with respect to the local field potential [Cassenaer and Laurent, 2007].
8.6.3 Neural Modelling of Combination

The current neural model is described at both neural and algorithmic levels. As discussed, non-directional inhibition is not likely to be amenable to a fully neural implementation. An attempt to produce a neural implementation of the non-directional combination scheme would help to assess its plausibility.

8.6.4 Adding Further Modalities

The optical flow experienced by a cricket during a turn to sound is related in a relatively simple way to certain other sensory inputs. For example no optical flow is received in the dark, and failing to take this into account might lead the robot, equipped with an efference copy, to turn too fast in the dark.

Input to the liquid state machine is inherently multidimensional, and could be multimodal. If an ambient light-level was fed into the liquid state machine in addition to the auditory efference copies then, assuming adequate training, the robot ought to be able to extract a rule in which the sensory prediction is dependent on the ambient light level. Similarly, the robot’s on-board program contained a simple rule to stop if the infra-red sensors detected a nearby obstacle. This also resulted in an absence of optical flow, while the LSM continued to receive auditory efference copies. In order to learn an appropriate sensory prediction in this case the sensor readings could also be inputs.

The liquid state machine architecture could also be used in the incorporation of new sensors, such as the wind sensors described in [Chapman and Webb, 1999]. Cercal wind sensing in the cricket is another modality, besides vision, which generates reafferent input during self motion.

8.7 Co-ordination of Behaviour

At the beginning of this dissertation it was claimed that the examination of interactions between behaviours ought to increase the chances of finding general mechanisms. The problem of dealing with sensory re-afference is certainly a general problem. What could the current work tell us about other animals?

It has still not been shown that insects use complex forward models. However, having investigated the five hypotheses of Chapter 2 the idea that a cricket employs an efference copy to predict its optomotor re-afference still stands. This implies the
mapping of discrete signals with fast dynamics from the auditory system into slow sustained signals matching the optomotor response. The robotic model has shown that this can be achieved by a biologically plausible recurrent network containing orders of magnitude fewer neurons than are contained in the mushroom bodies of the cricket brain. In the light of this work the possibility that even the simplest of insects could compute with efference copies is supported.
Appendix A

Phonotaxis on the Trackball

A large body of previously published data characterising phonotaxis was available to draw upon (reviewed in Section 2.4), but a small amount of new data was needed for the present study. Experiments on the trackball apparatus of Section 4.2.2 served two purposes: finding the correspondence between steering measurements on the trackball and in the arena, and looking for adaptation of the response over long periods of sound stimulation.

The experimental methods were as described in Section 4.2. For most phonotaxis tests the patterned cylinder used to provide visual stimuli (see section 4.2.3) was removed, and the speakers were placed at either 90° or 45° to the mid-line of the cricket. The fluorescent tube inside the sound-proof box was switched off in order to preclude any confounding effects from form discrimination. A total of 12 crickets were tested. Two of these crickets were also tested with the patterned cylinder in place, to check that it did not disrupt the sound field enough to abolish auditory orientation.

All phonotaxis experiments involved playing the calling song to the insects alternately from the left and right, switching every thirty seconds. Because phonotaxis is a reactive behaviour this alternating presentation meant that a responsive animal would undergo a rapid change in its steering direction shortly after each switch. A non-responsive cricket, on the other hand, might steer towards the side of the initially active speaker due to internal bias or uneven tethering, but would not change its steering when the angle of the sound changed.
A.1 Speakers at 90°

Once placed on the trackball crickets tended either to walk spontaneously or to walk when the calling song was started. Auditory orientation was assessed by looking at the left–right component of the trackball output. Crickets which steered strongly (like the top example shown in Figure A.1) turned continuously towards the active speaker, switching direction when the sound source was changed.

Not all crickets steered so strongly, and not all individuals steered to the same extent every time they were tested. In order to summarise the steering strength on a given trial the average steering speed towards the active speaker was calculated from the L–R displacement trace. For each thirty second period of sound the mean steering speed was read off as shown in Figure A.2. The overall mean for all left and right periods was taken to represent a bias of the cricket or the experimental set-up and was subtracted from the values for the individual periods. After this adjustment the mean of all the left periods was calculated. This is the average steering speed towards the active speaker, which is positive for positive phonotaxis and negative for negative phonotaxis.

When the distribution of steering speeds (Figure A.3) is examined there is a clear mode at 0.13 – 0.2 cm/s. The example shown at the top of figure A.1 has a much higher steering speed (0.58 cm/s) which is in the top 8% of the distribution. An example from a more typical trial, with mean steering speed 0.14 cm/s is shown at the bottom of Figure A.1.

Ideally it would be possible to decide whether a trial did or did not demonstrate phonotaxis based on this steering speed measurement. However Figure A.3 shows that this is not really feasible. If trials within the modal group of the histogram truly show phonotaxis, as Figure A.1 suggests they do, then what lower score should a trial receive before it is discounted? It is reasonable to say that the two trials with negative steering speeds did not demonstrate (positive) phonotaxis overall. However the problem with low values of either sign is that they could be a result of random meandering, a mixture of positive and negative phonotaxis and pauses, or just phonotaxis with a low overall level of activity.

More concretely, it is certainly possible to say that the distribution of steering speeds shows a significant bias towards positive phonotaxis (t-test, t=6.7, df=37, p=7.3e-8). In the multimodal experiments the tricky problem of classifying phonotaxis in individual cases has been avoided by looking at the effect of visual input on averages of populations of crickets. This contrasts with the approach taken by some
Figure A.1: Example responses to the standard phonotaxis test with speakers at 90°.
Top axes: Part of a trial with mean L–R steering speed of 0.58 cm/s. Bottom axes: Part of a trial with mean L–R steering speed of 0.14 cm/s. A slight bias to the left is evident in both trials.
Figure A.2: Measuring phonotactic steering from the L–R displacement graph. Example for one period of sound from the left. $dx$ — Change in displacement. $dt = 30$ s — Duration of one period of sound presentation. $dx/dt$ — Average steering speed. Left steering is nominally positive, right is nominally negative.

Figure A.3: Distribution of phonotactic steering strengths across all trials. 38 Trials from 12 crickets. Not all crickets contributed the same number of trials.
A.2 Speakers at 45°

The behaviour of crickets when the speakers were placed at 45° rather than 90° was qualitatively similar. Only six of the twelve insects were tested in both conditions. When the change in steering speeds is examined between the two conditions there is a trend for the crickets to have walked faster in the 45° case, as shown in Figure A.4. For the small number of insects tested is not possible to class this change as significant (Wilcoxon signed ranks test, T = 3, p = 0.16). This failure of the crickets to distinguish between the two stimulus angles does not suggest that sound localisation capabilities are likely to be poor; simply determining the correct side of the song would be enough to allow navigation towards a sound source with high meandering. Additionally, close-up recordings of phonotaxis in the arena (Appendix C) suggest that auditory acuity may be better in the frontal field between +45° and −45°.

![Figure A.4: Average steering speeds towards the active speaker for six crickets, with speakers at 45° and 90°.](image)

A.3 Response with cylinder in place

When the patterned cylinder used to create optical flow was placed in position during a phonotaxis test (with the sound-proof box kept dark) the two animals tested were
still observed to turn towards the active speaker. The average steering speed was reduced compared to their behaviour without the cylinder present. Further phonotaxis trials with the drum present and visible were conducted as part of the time-averaged multimodal tests (Section 4.4.1).

A.4 Maintenance of phonotaxis over several minutes

Most phonotaxis experiments were eight minutes long. As some of the multimodal experiments were this long or longer it is useful to know how the phonotactoc response changes over this time-scale. In particular it is necessary to know if it reduces as the cricket expends energy, becomes less motivated, or becomes habituated to the sensory input.

Bias-adjusted average steering speeds had already been calculated for each 30-second period of sound presentation. Figure A.5 shows the distribution of this quantity (ignoring the direction of turning) over the full eight minutes of the 90° trials.

![Figure A.5: Distribution of phonotaxis responses over eight minutes. Each box plot shows absolute steering speed for 12 crickets responding to calling song from 90°.](image)

Looking at the medians for each time period it appears that there is a reduction in turning over the first three minutes. Beyond three minutes the median turning speed stays roughly the same. In the portion of the graph where the median is reducing there is also a change in the distribution of the steering speed. Initially the inter-quartile range is larger, with a skew towards zero. When the median steering speed levels out the inter-quartile range has become narrower, and the skew is less apparent. This change in skew suggests that some crickets take a couple of minutes to become fully
responsive after the sound starts.

The graph shows that it is not safe to assume that phonotactic turning will be maintained at a constant level throughout a long trial. In particular there may be a reduction in average steering speeds over the first three minutes.
Appendix B

Scototaxis in the Arena

Prior to the arena experiments of Chapter 5 a number of trials were run to try and observe scototaxis behaviour from the crickets. The original aim was to look at choice behaviour similar to that described by [Stout et al., 1987]. Although it was possible to elicit scototaxis it turned out that the phonotaxis response dominated for the combination of visual target and calling song that was tried.

Since orientation towards dark stripes turned out to be relevant to the interpretation of the later experiments the results of those initial trials are presented here. They illustrate the kinds of walking paths produced when crickets orientate towards visual targets, and show that the distracting effect of those visual targets is not always sufficient to interfere with phonotaxis.

B.1 Scototaxis Without Calling Song

A small number of trials were conducted to test for orientation to dark shapes (scototaxis) in the absence of calling song. [Böhm et al., 1991] reported that a narrow vertical stripe was most effective in their experiments, so a black stripe 23 by 5 cm was made out of black cardboard and fixed to the arena wall. On each trial a cricket was placed directly into the centre of the arena with no cover and recording was immediately started using the overhead camera.

The 23 by 5 cm target was successful in eliciting scototaxis behaviour from the crickets, whether positioned in a vertical or a horizontal orientation. Figure B.1 shows four of the recordings obtained. Each example is from a different cricket.

A trial was considered to show scototaxis if the first path segment that reached the wall did so within the target’s boundaries (top-right example), or if the cricket
Figure B.1: Examples of walking paths showing scototaxis in the arena with a horizontal black target 23 by 5 cm. Each track is from a different cricket.
displayed repeated approaches to the target in preference to other regions of the arena perimeter. All four of the recordings shown were classed as showing scototaxis. Trials not classed as showing scototaxis mostly showed either inactivity or wall-following behavior. Over 24 trials the probability of scototaxis being seen was 0.63.

B.2 Influence of Attractive Targets on Phonotaxis

The black target used in the scototaxis trials had proved to be attractive to freely walking crickets. In a subsequent experiment the standard calling song was played to crickets from one of the concealed speakers, while a black stripe (32 by 13.5 cm) was placed vertically, either equidistant between the two speakers or alongside the active speaker. It was hypothesised that the presence of the target might alter the directness of the crickets' paths by distracting them from the speaker when placed centrally, or providing a stabilising cue when placed next to the speaker.

Eleven crickets were tested for phonotaxis in the blank arena, of which eight walked. In each of the two visual conditions seven of the eleven crickets walked, all of which had also walked in the blank condition. The tracks are shown in figure B.2 together with a graph of the associated D values.

Most tracks from the three conditions look similar, although there are a few unusual examples. In the condition with the visual target between the speakers one cricket appears to change course to approach the target. In the condition with the visual target beside the active speaker there is a track that has several sharp changes of course. These alone are, however, not enough to affect the mean D values which are very close: 0.293 (Phono); 0.291 (Equidistant); 0.326 (Coincident).

Appendix B. Scototaxis in the Arena

Figure B.2: Phonotaxis with and without attractive target. Top left — Blank Arena. Top right — Attractive target equidistant between two speakers. Bottom left — Attractive target beside active speaker (only one target present at a time). Bottom right — Values of D statistic obtained for each condition.
Appendix C

Close-Up Recordings in the Arena

In order to find out more details about the body orientation of *Gryllus bimaculatus* during phonotactic walking a small number of experiments were carried out in the arena of Section 5.2.1 in which a hand-held video camera was used to capture close-up film. The resulting data was used to provide descriptive statistics about the magnitude and speed of the animal's rotations during ordinary phonotaxis. These values were used to correctly scale motor parameters of the robot model.

C.1 Data Capture

The camera (Sony DCR-TRV series) was held at a distance of around 40 cm above the crickets and kept as vertical as possible. As a cricket walked the camera was moved with it to keep the cricket within the image. The video data was recorded at 25 frames per second with a resolution of 720 x 576 pixels. In order to use the hand-held camera it was necessary to open the door at one end of the arena, behind the speakers. The crickets would have been able to see the camera itself, the person holding the camera, and some of the room beyond the soundproof box.

During these close-up recordings the floor of the arena was covered in cardboard marked with a 5 cm square grid of small black dots as shown in Figure C.1. These acted as reference markers with respect to which the orientation of the cricket and the camera could be measured.
C.2 Tracking

Videos from the close-up recordings were processed using the same software described in Section 5.2.7. The lens of the camcorder did not significantly distort the useful part of the images so no correction step was required. In order to allow the orientation of a cricket to be calculated it was necessary to track two points on its body. The first was the front of the head, equidistant between the palps and antennae which were clearly visible from above. The other point chosen was the tip of the abdomen, where the ovipositor (and wings in long-winged morphs) protruded. Once these two points had been tracked the software automatically added a third point, equidistant between the other two.

The reference dots on the arena floor were also tracked throughout these sequences (see Figure C.1). In the first frame four reference dots surrounding the cricket were chosen. These were tracked until one or two became close to the edge of the video image, at which point the dots about to be lost were swapped one-by-one for new markers entering the image on the other side. This technique of swapping was continued throughout each sequence, ensuring that a 2-D frame of reference could always be traced back to the starting image. The tracking software contained a function which used this sequence to reconstruct the path of the cricket in arena co-ordinates rather than image co-ordinates. Figure C.2 shows two examples.

Figure C.1: Example close-up video frame showing the 5 cm floor grid with four dots marked, and the cricket with the head position marked.
C.3 Separating Behaviour into Bouts

Crickets tended to walk in bouts, reaching similar peak velocities during adjacent bouts, with distinct dips in speed or pauses in between. The largest changes in direction during a path occurred after pauses or dips, whereas only smooth adjustments were made during bouts. This made it interesting to separate the paths into bouts, to examine the contribution of each one separately. The speed of the middle point of the tracked cricket was used for this process.

Due largely to noise in the tracking process it was found that using a simple speed threshold to identify pauses and dips did not produce a qualitatively good result. This was particularly the case for overhead recordings, where the spatial resolution meant that occasional false minima were created. During tracking a cricket’s location was defined as the centre of mass of a group of pixels of a specified colour. This value could therefore have sub-pixel resolution. However lighting conditions in certain regions of the arena occasionally meant that only a small number of cricket-coloured pixels could be identified. In this case a slowly moving cricket might be identified by the same set of pixels in adjacent frames. This created false minima in the speed trace where a cricket appeared not to move for a single time step. To remove false minima and other noise, speed traces were first smoothed using a Gaussian kernel of full width 200 ms at half maximum. A time-step was then marked as being part of a dip or pause (the two were not explicitly distinguished) if its speed was less than one third of the maximum speed during the trace, and it was part of a local minimum or extended pause. These were defined as sequences of time-steps consisting of a deceleration of greater than 3.75 cm/s\(^2\), zero or more time-steps with fluctuations no greater than this, and an acceleration of greater than 3.75 cm/s\(^2\). These values were chosen to give qualitatively reasonable segmentation.

C.4 Results

Ten close-up recordings were captured in the arena with no additions to the environment described in Section 5.2.1 except the patterned floor covering described in Section C.1. Two example traces are shown in Figure C.2. The properties of these and the other eight traces were combined and are presented below.
Figure C.2: Tracker output for two paths recorded using the hand-held camera. The 5 cm floor grid has been drawn on with lines although the real grid was marked only by dots (see Figure C.1). Each path consists of three separate lines indicating the course of the head, abdomen and middle point. When the crickets walk straight or in smooth curves the three lines are difficult to separate visually, but when sharp turns occur (two examples marked 'X', left) there is some sideways movement and the head in particular can be seen to follow a different course. The right-hand example shows distortion of the floor grid towards the end of the track, caused by rotation of the camera.
A Note on the Presentation of Path Properties

Due to the relatively small number of crickets involved in some tests path property distributions are displayed as box plots rather than histograms. In these diagrams the central box encompasses the inter-quartile range \((Q_1 - Q_3)\), and the whiskers encompass points without the inter-quartile range up to a limit of \(\frac{3}{2} \times (Q_3 - Q_1)\). Points outside this limit are marked individually.

Path Speed

The median (6.23 cm/s) and spread of the mean path speeds are shown in Figure C.3. The values are rather evenly distributed, which is reflected in the closeness of the mean (6.25 cm/s) to the median.

![Box plot showing median, range and inter-quartile range of mean path speeds for ten crickets in the blank arena.](image)

Instantaneous Speed

All crickets made a series of spontaneous pauses or dips in speed (hereafter referred to jointly as pauses) on the way to the speaker. In seven of the recordings the crickets also made pauses upon reaching a join in the floor covering, artificially shortening the preceding walking bouts. The associated bouts were excluded from the analysis. In the final two recordings the floor covering was re-orientated to eliminate this behaviour.

Examining the mean interval between pauses gives us the periodicity of the walking bout pattern. The distribution for the ten crickets recorded is shown in Figure C.4.

The median pause interval between crickets is 1.12 s, corresponding to a walking bout frequency of 0.9 s\(^{-1}\). Two crickets which had particularly sparse pauses contribute to a slightly higher mean of 1.33 s.

When the length of the pauses is taken into account the mean duration of running bouts between crickets is 1.10 seconds, with a mean speed of 6.50 cm/s. The actual
mean bout distance between crickets is slightly higher than would be predicted from these values, at 7.99 cm.

As mentioned in Section C.3 the crickets tended to reach similar peak velocities in adjacent bouts. The mean peak speed over all bouts was calculated for each cricket. Figure C.5 shows the distribution.

**Figure C.5:** Distribution of peak instantaneous speed. Means of peak bout speeds in all bouts for 10 crickets.

## Turning Behaviour During Running Bouts

When *Gryllus bimaculatus* were tested on an open loop trackball system the rotational turning tendency was found to be related to the angular position of the active speaker in a sinusoidal manner [Böhm et al., 1991]. Turning was maximised when the speaker was at 90° to either side, and decreased as the speaker was moved to the front or rear. The tuning curves shown in [Böhm et al., 1991] are long-term time averages, so it is not possible to say exactly how the walking pattern alters to effect this change. For example the crickets might increase their maximum speed, or they might make fewer pauses. To test for the equivalent behaviour in the arena the angle rotated during each walking bout was plotted against the speaker angle during the preceding pause (Figure C.6, left).
C.4. Results

Figure C.6: Left – Body rotation during each walking bout vs. speaker angle during the previous pause. 73 walking bouts from 10 crickets. Right – As left but with points where $x < -90^\circ$ reflected in $x = -90$, and regression line superimposed. Correlation coefficient $r = 0.56$ which is significant at the 5% level.

The first thing to note about the resulting graph is that the crickets have not experienced the speaker at all possible angles. The speaker is within $45^\circ$ of the mid-line on 78% of stops. Only six bouts followed a pause where the speaker was positioned caudally, and these are clustered together. This makes it difficult to say what the overall shape of the curve would be. Within the frontal field, however, it appears that there is a linear relationship between the two variables. This is tested by calculating the correlation coefficient and linear regression, shown in Figure C.6, right. Because the physical tympanal response [Michelsen, 1998] is symmetrical front-to-back the caudal points have not been rejected, but reflected in the line $x = -90$. The correlation coefficient is 0.56 to 2 d.p. which is significant at the 5% level.

Having established this relationship it is reasonable to ask how the cricket alters the amount of rotation during a bout. It could be a case of changing the average rate of rotation throughout the bout, but since bout lengths are not fixed the rate of rotation could feasibly stay constant, and the cricket might simply walk further, or faster. Examining the relationship with each of these parameters in turn shows that the first explanation is the correct one. Figure C.7 shows the scatter graph for each variable against the speaker angle during the pause with the associated regression line and correlation coefficient. Only the rotation speed shows a significant correlation with speaker angle at the 5% level (correlation coefficient 0.54).
Figure C.7: Top Left – Rotation speed during each walking bout vs. speaker angle during the previous pause. $r = 0.54$, significant at 5% level. 73 walking bouts from 10 crickets. Top Right – Bout duration vs. speaker angle, $r = 0.15$ (NS). Bottom Left – Mean speed over bout vs. speaker angle, $r = 0.17$ (NS).
C.5 Conclusions

Close-up recording of crickets in the blank arena confirmed the observation from Chapter 5 that the paths of the crickets tended to show less deviation than was predicted. The effect may have been enhanced because the arena door had to be opened to admit the camera, introducing further visual contrast in the room beyond.

The most important conclusions for the robotic model are that the pauses made by the crickets do not seem to be correlated with the chirp pattern, and that the speed of turning is correlated to the lateral deviation of the speaker within the frontal field.
Appendix D

Diagram For Auditory Board

Drawn by Rob Macgregor
Figure D.1: Diagram for Auditory Board
Appendix E

Parameters For Predictive Model

The parameters of the liquid state machine were largely the same as those used by [Maass and Natschläger, 2002] and are listed below. Differences between the current implementation and that of Maass and Natschläger are highlighted at the end of this section.

Neurons

Neurons of the recurrent pool were of the leaky integrate-and-fire type (Section 6.5.2.1) with the following parameter values: $C = 30$ (nF); $R = 1$ (MΩ); $v_{\text{rest}} = -60$ (mV); $v_{\text{thresh}} = -45$ (mV); $v_{\text{reset}} = -46.5$ (mV); $I_{\text{ext}} = 13.5$ (nA). The initial values of $v_m$ were uniformly distributed between $v_{\text{reset}}$ and $v_{\text{thresh}}$. Conceptually the neurons were organised in a 3-D lattice of $3 \times 3 \times 15$ units. 20% were chosen to be inhibitory at random. Inhibitory neurons had a refractory period of 2 ms and an axonal delay of 1 ms. Excitatory neurons had a refractory period of 3 ms and an axonal delay of 2 ms.

Connectivity

Connections between the neurons of the 3-D lattice were assigned randomly. The probability of a connection being made between a neuron $a$ and any other neuron $b$ on the lattice decreased with distance according to

$$P(\text{connect}(a, b)) = C \cdot e^{-\frac{(D(a,b)/\lambda)^2}{2}}$$

where $D$ is the Euclidean distance between $a$ and $b$, and $\lambda$ is a parameter that can be used to alter the density and average distance of connections. For the current work $\lambda$
was always set equal to two. Self connections were not allowed. The constant $C$ varied depending on the type of the pre- and post-synaptic neuron as detailed in Table E.1.

<table>
<thead>
<tr>
<th>Pre-synaptic Type</th>
<th>Post-synaptic Type</th>
<th>$C$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excitatory</td>
<td>Excitatory</td>
<td>0.3</td>
</tr>
<tr>
<td>Excitatory</td>
<td>Inhibitory</td>
<td>0.2</td>
</tr>
<tr>
<td>Inhibitory</td>
<td>Excitatory</td>
<td>0.4</td>
</tr>
<tr>
<td>Inhibitory</td>
<td>Inhibitory</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Table E.1: Dependency of $C$ on pre- and post- synaptic neuron types

**Synaptic Parameters**

All synapses between neurons of the recurrent pool were full dynamic synapses (Section 6.5.3.4). The behaviour of a synapse is determined by five parameters: the intrinsic weight, $A$ (nA); the time constant of the synaptic current, $\tau$ (ms); the base rate of depression, $U$; the time constant of depression, $\tau^{dep}$ (ms); the time constant of facilitation, $\tau^{facil}$ (ms). $\tau$ was 3 ms for excitatory synapses and 6 ms for inhibitory synapses. The other parameters were randomly sampled from Gaussian distributions whose means depended on the types of the pre- and post-synaptic neurons as detailed in Table E.2.

<table>
<thead>
<tr>
<th>Pre-synaptic Type</th>
<th>Post-synaptic Type</th>
<th>$A$ (nA)</th>
<th>$U$</th>
<th>$\tau^{dep}$ (ms)</th>
<th>$\tau^{facil}$ (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excitatory</td>
<td>Excitatory</td>
<td>30</td>
<td>0.5</td>
<td>1100</td>
<td>50</td>
</tr>
<tr>
<td>Excitatory</td>
<td>Inhibitory</td>
<td>60</td>
<td>0.05</td>
<td>125</td>
<td>1200</td>
</tr>
<tr>
<td>Inhibitory</td>
<td>Excitatory</td>
<td>-19</td>
<td>0.25</td>
<td>700</td>
<td>20</td>
</tr>
<tr>
<td>Inhibitory</td>
<td>Inhibitory</td>
<td>-19</td>
<td>0.32</td>
<td>144</td>
<td>60</td>
</tr>
</tbody>
</table>

Table E.2: Dependency of synapse parameters on pre- and post- synaptic neuron types

For $A$ the standard deviation of the distribution was 100% of the mean, and for all other parameters it was 50% of the mean. When sampling values for $A$, a synaptic weight of the wrong sign was replaced by a weight of zero. This would happen to 15.8% of synapses on average. Similarly, $U$ was required to be greater than or equal to zero. Negative values were replaced with zero, meaning that that synapse would
not undergo any depression (average incidence 2.2%). $\tau^{dep}$ and $\tau^{facil}$ were required to be greater than zero, and negative values were replaced by a time constant of 1 ms (average incidence 2.2%).

**Inputs to the Recurrent Pool**

Inputs to the recurrent pool came from the spike trains of the auditory circuit’s BN neurons via a pair of axons adjusted to approximate the lag in the robot’s optomotor system. For a few tests detailed in sections 7.5.2.3 and 7.5.2.4 input was also taken from the optomotor system via two spike trains, one for each direction of motion. Each input spike train made synapses onto a randomly chosen 30% of the neurons of the recurrent pool. The synapses were of the simple exponential type (Section 6.5.3.3), with a time constant ($\tau$) of 3 ms. Synaptic weights ($A$) were selected randomly from Gaussian distributions with means 18 nA for projections onto excitatory neurons and 9 nA for projections onto inhibitory neurons. As for connections within the recurrent pool the standard deviations of the Gaussian distributions were 100% of their means, and negative values were replaced with zero.

**Output Filters**

The output spike train from each neuron of the recurrent pool was low-pass filtered by a leaky integrator with time constant 30 ms, to simulate the effect of the spikes on the membrane potential of a post-synaptic neuron. These filtered signals, rather than the spike trains, were the input to the readout unit (Section 6.5.6).

**Differences From the Model of**

[Maass and Natschläger, 2002]

**Resting Neural Membrane Potential**

Maass and Natschläger give $v_{rest}$, $v_{thresh}$ and $v_{reset}$ values of 0, 15 and 13.5 millivolts respectively. This makes no difference to the dynamics of the model neurons. The reason for using values within a biologically plausible range in the robot’s controller was so that the neurons could use conductance-based synapses for input if necessary.
Axonal Delays

Maass specifies axonal delays of 1.5 ms for excitatory–excitatory connections, and 0.8 otherwise. In the current model the implementation of axonal delays (using a queue) did not permit a resolution of less than one millisecond. Additionally, the axon length of a neuron was specified at its construction time, whereas post-synaptic targets were assigned separately at a later stage. This meant that delays could not be assigned on the basis of target type, and were instead dependent only on the type of the pre-synaptic neuron. Maass’s original specification was presumably based on measurements from mammalian cortex, and the change was not detrimental to the performance of the recurrent pool.

Synaptic Parameters

The values of all synaptic parameters were sampled from Gaussian distributions in the current work. In Appendix B of [Maass and Natschlager, 2002] it is stated that the intrinsic weight parameter $A$, was sampled from a gamma distribution, although in the body text (p. 2537) it is stated that a Gaussian distribution was used for the input synapses.

Maass and Natschlager also used a different strategy for dealing with sampled values outside the allowable range, stating that they were replaced with values from an appropriate uniform distribution. No further details were given.
Bibliography


