THE NERVOUS SYSTEM OF *CAENORHABDITIS ELEGANS:* A STOCHASTIC APPROACH

by

Bernie C. Till

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE (Computing Science)
in the School
of
Computing Science

© Bernie C. Till 1990

Simon Fraser University

October 1990

All rights reserved. This work may not be reproduced
in whole or in part, by photocopy or other means,
without the permission of the author.

APPROVAL

Name: Berndt Christian Till

Degree: Master of Science

Title of Thesis: The Nervous System of *Caenorhabditis elegans:* A Stochastic Approach

Examining Committee:

Chairman: Dr. Fred Popovich

Dr. James P. Delgrande
Associate Professor
Department of Computing Science
Senior Supervisor

Dr. Robert F. Hadley
Assistant Professor
Department of Computing Science

Dr. David L. Baillie
Professor
Department of Biological Science

Dr. Eduard Moser
Visiting Professor
Department of Computing Science
External Examiner

Date Approved: 27 FEB 1991

PARTIAL COPYRIGHT LICENSE

I hereby grant to Simon Fraser University the right to lend my thesis, project or extended essay (the title of which is shown below) to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users. I further agree that permission for multiple copying of this work for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this work for financial gain shall not be allowed without my written permission.

Title of Thesis/Project/Extended Essay

The Nervous System of Caenorhabditis Elegans: A Stochastic Approach.

Author:

(signature)

Berndt Christian Till

(name)

01 FEB 1990

(date)

ABSTRACT

The first nervous system to have its complete structure mapped is that of the soil nematode *Caenorhabditis elegans*. This thesis presents a stochastic theory of computation based on statistical mechanics and information theory and views the nervous system of *C. elegans* in its light. The underlying paradigm is that information exerts a force and that the dissipative dynamics of a nonlinear system are equivalent to information processing. The paradigm is developed in the context of the sensory preprocessing system and the central nervous system of the organism. Finally, a detailed model of the motor nervous system is presented using these techniques. The theoretical results are found to agree quite well with simulation results, which proves that the paradigm provides a fruitful theoretical framework.

ACKNOWLEDGEMENTS

The work reported herein would not have been possible without the raw data on the structure of the nervous system of *C. elegans* obtained by Dr. John White and his team at the Medical Research Council Laboratory of Molecular Biology, Cambridge England. It is with admiration and gratitude that this thesis is dedicated to him. Thanks are also due to those at Simon Fraser University whose vision, support and advice were instrumental in my own transition from amateur to scientist. Dr. Dave Baillie, who introduced me to Dr. White, and Dr. Peter Belton were among the earliest supporters of the work, and are responsible for my education in neurobiology. I am indebted to Dr. Tiko Kameda, without whom the work would never have begun, and to Dr. Joe Peters, whose enthusiasm gave me courage to proceed. I am also very grateful to Dr. Jim Delgrande, whose guidance and support were so vital to the project. Finally, I would like to thank Ed Levinson for many insightful and stimulating discussions.

Table of Contents

Introduction 1

The Non-equilibrium Boltzmann Machine 15
Modelling the Nervous System 21

The Structure and Physiology of *C. elegans* 30

Motor Postprocessing: The Noise Driven Filter 37
Mechanical Model of the Organism 55
Modelling the Nerve Cords 60
Time Evolution of the Means 68
Time Evolution of the Variations 70
Overall Time Response of One Section 76
Overall Time Response of the Nerve Cords 79

Summary 86

Discussion 91

Appendices

A - Sensory Preprocessing: Orthogonalization 97
Modelling the Sensory Preprocessing System 103

B - The Nerve Ring: Context Sensitive Association . . 116
Modelling the Nerve Ring 117
Time Response of the Nerve Ring 125
Characterization of the Impulse Response 130

C - Characterization of Travelling Waves 146

D - Solving for the Means and Variations 151
Time Evolution of the Means 153
Time Evolution of the Variations 156

References 166

Glossary of Terms 171

Index 183

LIST OF FIGURES

Figure 2.1 - The main anatomical features of *Caenorhabditis elegans* . 31
Figure 5.1 - Spatial organisation of the motor nervous system and somatic musculature 40
Figure 5.2 a - Connectivity of the motor nervous system and somatic musculature 41
Figure 5.2 b - Connections reported by White, et al 42
Figure 5.2 c - Extrapolated connections 43
Figure 5.3 a - Connections within the dorsal and ventral nerve cords . 44
Figure 5.3 b - Connections between the dorsal and ventral nerve cords 45
Figure 5.3 c - Neuromuscular junctions (extrapolated) 46
Figure 5.4 a - Connections of the DA motor neurons 47
Figure 5.4 b - Connections of the DB motor neurons 48
Figure 5.4 c - Connections of the DD motor neurons 49
Figure 5.4 d - Connections of the VA motor neurons 50
Figure 5.4 e - Connections of the VB motor neurons 51
Figure 5.4 f - Connections of the VD motor neurons 52
Figure 5.5 - Idealized structure of one section 53
Figure 5.6 - Approximate posture of the body 56
Figure 5.7 - Section geometry and force balance 58
Figure 5.8 - The subsystems of a single section 62
Figure 5.9 - The first three sections, showing inter-section coupling 63
Figure 5.10 a - Mean time response of the DD and VD neurons 69
Figure 5.10 b - Mean time response of the DB and VB neurons 69
Figure 5.10 c - Mean time response of the muscular excitations . . . 69
Figure 5.11 a - Differential time response of the DD and VD neurons . 73
Figure 5.11 b - Differential time response of the DB and VB neurons . 73
Figure 5.11 c - Differential time response of the muscular excitations 73
Figure 5.11 d - Time evolution of the differential muscle lengths . . 73
Figure 5.12 a - Time evolution of x_{DV} in the linear model 74
Figure 5.12 b - Time evolution of x_{BV} in the linear model 74
Figure 5.12 c - Time evolution of x_{EV} in the linear model 74
Figure 5.12 d - Time evolution of x_L in the linear model 74
Figure 5.13 a - Differential time response of the DD and VD neurons . 75

Figure 5.13 b - Differential time response of the DB and VB neurons . 75
Figure 5.13 c - Differential time response of the muscular excitations 75
Figure 5.13 d - Time evolution of the differential muscle lengths . . 75
Figure 5.14 a - Overall time response of the DD and VD neurons . . . 77
Figure 5.14 b - Overall time response of the DB and VB neurons . . . 77
Figure 5.14 c - Overall time response of the muscular excitations . . 77
Figure 5.14 d - Time evolution of the differential muscle lengths . . 77
Figure 5.15 a - Overall time response of the DD and VD neurons . . . 78
Figure 5.15 b - Overall time response of the DB and VB neurons . . . 78
Figure 5.15 c - Overall time response of the muscular excitations . . 78
Figure 5.15 d - Time evolution of the differential muscle lengths . . 78
Figure 5.16 - Differential muscle lengths for coupled sections . . . 83
Figure B.1 a - The sigmoid transfer function of a neuron 123
Figure B.1 b - The derivative of the sigmoid 123
Figure B.1 c - The second derivative of the sigmoid 123
Figure B.2 a - The first six decreasing nonoscillatory eigenfunctions 138
Figure B.2 b - The effect of the sigmoid on these functions 138
Figure B.3 a - The first six increasing nonoscillatory eigenfunctions 139
Figure B.3 b - The effect of the sigmoid on these functions 139
Figure B.4 a - Decreasing linear oscillatory eigenfunctions of index 0 140
Figure B.4 b - Decreasing linear oscillatory eigenfunctions of index 1 140
Figure B.4 c - Decreasing linear oscillatory eigenfunctions of index 2 140
Figure B.5 a - Decreasing nonlinear oscillatory eigenfunctions, index 0 141
Figure B.5 b - Decreasing nonlinear oscillatory eigenfunctions, index 1 141
Figure B.5 c - Decreasing nonlinear oscillatory eigenfuncions, index 2 141
Figure B.6 a - Increasing linear oscillatory eigenfunctions of index 0 142
Figure B.6 b - Increasing linear oscillatory eigenfunctions of index 1 142
Figure B.6 c - Increasing linear oscillatory eigenfunctions of index 2 142
Figure B.7 a - Increasing nonlinear oscillatory eigenfunctions, index 0 143
Figure B.7 b - Increasing nonlinear oscillatory eigenfunctions, index 1 143
Figure B.7 c - Increasing nonlinear oscillatory eigenfunctions, index 2 143
Figure C.1 - Wave propagation due to coupled oscillations 146
Figure C.2 - A travelling wave at successive instants 147

INTRODUCTION

What is intelligence? This simple question has come to preoccupy many of our best thinkers. Enormous resources are devoted to the achievement of artificial intelligence. Yet, progress has been considerably slower than expected. The results obtained by the automation of formal logic and other axiomatic systems, while significant in their own right, have been disappointing in so far as they have demonstrated that the design of intelligent machines is a much more difficult task than originally anticipated.

An intelligent entity should have at least the following abilities: the ability to recognize a problem when it encounters one whether or not it belongs to a known class of problems, the ability to decompose it into its component parts in such a way as to facilitate solution, the ability to recognize applicable methods of solution and to develop new methods when necessary, and finally, the ability to recognize an acceptable solution when one is found. Ideally, all this should be possible regardless of the domain of discourse.

This is not to say that these abilities are equivalent to intelligence, but that a system which does not possess them is certainly not intelligent. It remains an open question whether these abilities, when carried to a sufficient degree of complexity, are sufficient to account for the subjective experiences which, for example, human beings tend to equate with being intelligent. In this respect, it is quite

difficult to resist the temptation to read more into words like intelligence and cognition than is strictly permitted by their usage within any well defined paradigm of information processing.

As a result, it is important to distinguish between the internal processes which take place when a human being engages consciously in deliberately structured thought processes which correspond to the abilities described above, and the internal processes which occur when a human being relinquishes conscious control of his thoughts and simply allows these abilities to operate autonomously. The first case is an example of the application of learned behaviour, while the second case is an example of exercising the naturally occurring cognitive faculties. At no point should any remark in this thesis be construed as referring to the first case.

Whether these two cases are actually different in any fundamental way is still very much an open question, but the postulate that they are fundamentally different is one of the conjectures which motivate the formulation of the paradigm presented below. While this conjecture is not fundamental to the theoretical machinery developed below and is not proven thereby, it is hoped that this machinery will facilitate further work which sheds light on the matter.

In any event, the above abilities are certainly beyond the scope of any artificial system currently existing. Many, if not most, of these systems are grounded in the paradigm of abstract symbol manipulation. The approach begins by asking

what types of symbols and symbol manipulation primitives are needed, how these may be combined to form more complex structures and what kinds of algorithms are required in order to build effective systems layered on these elements. It proceeds by implementing such a system and measuring its performance. Then these measurements are analyzed in the hope of refining the symbols and algorithms in order to improve the performance of the system. The underlying assumption is that, if the right set of symbols and algorithms could be discovered, then their implementation would result in an intelligent system.

This approach has fallen short of the desired goal often enough that it seems necessary to question the paradigm. The precise nature of basic difficulties with the paradigm is a matter which is surrounded by deep controversy, but a number of issues repeatedly come to the surface. Of significance in the context of the work presented here are the limited expressive power of abstract symbols, the fundamentally deductive nature of the primitives, and the inherent obstacles to the consistent representation of uncertainty. To implement a system within this paradigm, one must predefine an arbitrary mapping between the form of the symbols and their meaning, one must simulate induction in a deductive framework, and one must simulate uncertain reasoning with reasoning about uncertainty.

A fruitful line of inquiry might thus be to investigate systems which are not subject to these limitations. This would require that the symbols have the property that form is

function. It would require the processing to be fundamentally inductive and to arise as a direct result of the coexistence of the symbols. And it would require the representation and processing of uncertainty to be inherent in the representation and processing of information.

Such a line of inquiry would demand a fundamental re-examination of the notions of what constitutes a problem and a solution and how to construct a system that arrives at the latter when presented with the former. Taking into account the requirements of the previous paragraph, one is led to think less in terms of finding solutions to problems, and more in terms of transforming problems into solutions. This, in turn, leads to a re-examination of the very nature of information and information processing. One is faced, in short, with the task for formulating an alternative paradigm, beginning with first principles.

In so doing, one finds oneself on theoretical ground which is many levels removed from considerations such as the nature of consciousness or the subjective experience of being an intelligent entity. An alternative paradigm, along the lines which are beginning to emerge, would necessarily have nothing to say about such matters. The primary motive for constructing such a paradigm is simply the hope that it will, when sufficiently developed, provide a viewpoint and a theoretical framework which allow such high level issues to be approached more effectively than otherwise. For the present, however, we must be content to observe nature and lay the

foundations of the paradigm.

In nature, we find that solutions to some problems are found in a very different manner from that which is embodied in the paradigm of abstract symbol manipulation. Solutions come into being simply as a result of the existence of a problem. The working hypothesis underlying the present work is that this process is called evolution or cognition, according to the time scale and level of abstraction at which it occurs. One is led to ask, then, in what kind of system can any required symbols and primitives arise spontaneously, merely because they are required? The answer comes from the basic common ground that underlies the closely related fields of epistemology [Cox, 78], thermodynamics [Wannier, 66] and information theory [Jaynes, 57], [Jaynes, 78].

On this view, information is fundamental. In fact, there seem to be good reasons to believe that the mere existence of information is a sufficient condition for there to arise structures to process it. This is because information is energy. That is to say, energy is required to represent information and therefore a nonuniform distribution of information in a system is equivalent to a nonuniform distribution of energy. This gives rise to forces, which, in turn, cause a change in the configuration of the system, resulting a redistribution of the energy. If this change is irreversible, then the information content of the system has changed. In other words, information has been processed.

The structure of the system determines the forces which

result from a nonuniform energy distribution as well as the manner in which the system can change its configuration in response to forces. It also determines the conditions under which such changes will be irreversible. Thus, the structure of the system determines its capacity to process information. In some systems, the structure itself changes as a result of the response of the system to the forces. Thus, the capacity of the system to process information is altered by the act of processing information. This extremely important effect is called self organisation [Nicolis & Prigogine, 77]. It is observed in a wide variety of systems, from planetary systems to ecosystems to nervous systems [Harth, 82]. Its great significance for the proposed paradigm is that it has given rise to a body of theory which can be brought to bear on the principled interaction between procedural and declarative aspects of systems.

In this paradigm, an initial nonuniform distribution of energy, or information, is viewed as a problem presented for solution, the behaviour of the system in response is viewed as the process of solving the problem, and the final, usually more uniform, distribution of energy is viewed as the solution of the problem. As a simple example, consider a volume of water bounded by planes above and below, and let a temperature gradient be applied perpendicular to these planes. The system, composed of a large number of simple identical elements, namely water molecules, will organize itself into a pattern of convection currents. Hence, the problem of the

most efficient way to transport heat across the water volume has been solved. The solution was driven by the problem itself, with the help of a small amount of thermal noise.

This view of information processing represents a radical departure from the main stream in computing science. It forms the philosophical foundation of this thesis, which is motivated by a desire to understand the nature of information processing in a self organising dynamical system. This thesis argues that the functioning of the nervous system can best be understood in such terms, with the introduction of a level of indirection. That is, the nervous system provides a substrate in which a structure representing a problem can form and in which that structure can organize itself into a representation of a solution. The recognition of a problem and its solution arise out of the information being processed, under the influence of noise. When the uncertainty in these structures is significantly greater than the noise in the information, hitherto hidden degrees of freedom emerge to drive the processing. On the other hand, a signal so weak as to be buried in noise is ignored.

This is called maximum entropy inference, and is not only applicable to both deductive and inductive problem solving, but automatically applies the optimal blend of the two. Since it is driven by the problem itself, it is not limited by any given set of axioms and its referents are ultimately those features of the problem and solution domain which are self evident. Further, only those elements of that domain which

are relevant come into play. In contrast to the formal axiomatic system, all of this follows directly from the nature of the self organising system. Intelligent behaviour is an emergent property, arising from cooperation and competition between a large number of relatively simple, roughly identical elements. In this respect, the emergence of increasingly powerful neural structures by the cooperative/competitive interactions of individual organisms and the emergence of solutions by the cooperative/competitive interactions of representational states within particular neural structures are formally equivalent.

This paradigm follows from the working hypothesis, which remains unproven. In fact, both the hypothesis and the paradigm which follows from it have been called into question by prominent cognitive scientists [Fodor & Pylyshyn, 88]. Now, while the results obtained in this thesis do not address the ongoing controversy regarding this issue, they do elucidate certain mechanisms which can be called upon to support self organizing dynamics. The dynamics resulting from these mechanisms will likely be approximately equivalent, at least in some cases and under certain conditions, to those obtainable by clever programming, heuristics or knowledge engineering. Again, it is one of the unproven conjectures which form the philosophical foundation of this thesis that there will also result dynamical regimes which process information in ways that are quite different from, and more powerful than, a formal axiomatic system.

On this view, the deterministic behaviour of an axiomatic system is seen to be a limiting case of the stochastic dynamics described above. Specifically, it is the low temperature limit. When random fluctuations are superimposed on an information carrying signal, the resultant signal can be described in terms of an equivalent noise temperature. This allows one to bring the elegant and powerful formalisms of thermodynamics and statistical mechanics to bear at certain crucial points in the analysis of a system acted on by such a signal. On the other hand, as the energy content of these fluctuations approaches zero, the dynamics of the system will approximate more and more closely the rigidly algorithmic manipulation of a fixed set of abstract symbols.

As different as this paradigm appears, then, from that which underlies axiomatic systems, a clear place is reserved for the latter within the more general framework provided by the former. Thus, not only may one carry over all of the significant results so far obtained, but the paradigm allows one to reason about the circumstances under which these results apply. In addition, this approach clearly points the way for extending these results when the assumptions leading to strict determinism do not hold – that is, when the equivalent noise temperature of the signals is nonzero.

Even among the most cogent, carefully argued criticisms of the paradigm presented above, this issue is addressed either inadequately or not at all. For example, while the arguments presented by Fodor and Pylyshyn make an irrefutable

case for structure dependence in cognitive systems, it is difficult to understand why they claim that such mechanisms as energy minimization / entropy maximization and associative mechanisms in general are not structure dependent. More significantly, such arguments generally assume that, in the approach which underlies this thesis, the introduction of stochastic dynamics is intended to provide noise tolerance and permit algorithmic processes to be robust in the presence of noise. This could not be farther from the truth. In actuality, the intent is to provide a more general dynamics of information processing, which includes algorithmic behaviour as a limiting case.

This approach has yet another advantage: there already exist systems in nature which exhibit all of the properties mentioned. As a result, it is not necessary to work in a vacuum — one can begin by examining the nervous systems of living organisms. The paradigm described above provides a powerful framework in which to do this, in that it focuses attention on the relevant issues. Applying the paradigm to the analysis of a simple nervous system will indicate very quickly whether the paradigm is worth pursuing or not. The results obtained in this thesis show that the paradigm does, at least, provide an effective theoretical framework for this type of analysis. As well, they indicate promising directions for further inquiry.

Nervous systems can be classified into three levels of complexity, depending on how many stages of processing are

performed. The simplest are single stage systems. In such systems, there is no distinction between sensory and motor neurons, each one being driven by sensory input and driving a muscle or gland. In a two stage nervous system, sensory neurons synapse directly onto motor neurons. Finally, the three stage systems are characterized by the appearance of a plexus of interneurons. Even a very simple three stage nervous system has all of the functional elements that are required for the type of behaviour one would classify as intelligent when it reaches a sufficient degree of complexity.

Nonetheless, even simple nervous systems are formidably complex structures with numerous idiosyncrasies that can be classed as implementation details. It is therefore fortunate that the complete structure of a very simple nervous system, that of the soil nematode *Caenorhabditis elegans*, has recently become known. Using it as a prototype, this thesis argues that the functionality of the nervous system is best approached by adopting a three tiered paradigm: an adaptive, feature extracting preprocessing step, an associative step performing maximum entropy inference, and a relatively deterministic postprocessing step.

First, a stochastic model of computation is introduced in abstract terms. This is the non-equilibrium Boltzmann machine, which has formal properties amenable to analysis and which will serve as a theoretical point of departure. This computational model differs from that of Ackley, Hinton and Sejnowski [Ackley, Hinton and Sejnowski, 85] in two

significant respects. Firstly, it permits asymmetric link weights and thereby introduces non-trivial time dependent behaviour. The equilibrium states of the system are still determined by the symmetric part of the coupling matrix, but the path through state space followed by the system as it approaches these states is determined by the antisymmetric part. Secondly, it relaxes the assumption that the system has a single global temperature. Instead, it allows each node to be at a different temperature — namely the equivalent noise temperature its input signal. This results in a principled interaction between information and uncertainty, in that the midpoint gain of each node is determined by the noise on its input signal. The resulting dynamical regime is sufficiently general to incorporate, as special cases, all of the functional blocks of the nervous system of the organism.

The main body of the thesis, following a short section on the anatomy and physiology of the organism, examines the motor postprocessing system from a functional point of view, using techniques developed in the appendices in the context of the sensory preprocessing system and the central nervous system or nerve ring. The thesis focuses on the motor postprocessing system because of the availability of neurophysiological results against which the analytical results can be checked. Within the limited context of motor postprocessing, the results are in agreement with neurophysiological observations.

Experimental data are, unfortunately, much sparser for the other two functional blocks. A sufficiently detailed

cross check of the results must therefore await the availability of such data. To this extent, the analytical results obtained in the appendices must be regarded as speculative. It is to be hoped, however, that the availability of these results will stimulate biologists to conduct experiments which will furnish the data which will allow the predictions of the theory to be checked.

Nonetheless, these analyses of the three functional blocks form a theoretically coherent whole. The analytical viewpoints and methods developed in the discussion of the nerve ring, in particular, lay important groundwork for the analysis of the motor postprocessing system. The reader who is unfamiliar with the theory of skeleton filters, or who is interested in how such filters appear in the framework of a neural model based on the non-equilibrium Boltzmann machine, may thus wish to peruse the appendices before proceeding to the analysis of the motor postprocessing system.

The formulation of a well defined mathematical model of the motor postprocessing system required the development of a mechanical model of the balance of somatic forces associated with the movement of the organism. This model, being more comprehensive than those previously presented [Lee, 84], permitted the motor postprocessing system to be expressed in terms of a single system of differential equations. This is a striking example of the flexibility of the theoretical tools provided by the paradigm, which allowed both mechanical and neural coupling to be treated uniformly within a single

analytical framework.

Finally, the approach adopted in this thesis permitted Chalfie and White's model of the motor postprocessing system [Chalfie and White, 88], to be resolved into a system of gated, cross coupled oscillators and allowed forced oscillations to be ruled out quite definitely as a mode of operation for the system. Further, the approach allowed a detailed explanation of the gating mechanism, in terms of noise determined variable coupling, to be given.

In the last analysis, this thesis does not prove that the paradigm based on the idea that information exerts a force is superior to other paradigms. It does, however prove that the paradigm provides both a viewpoint and a set of analytical techniques that allow hard results to be obtained, at least in the case of simple neural structures. It also indicates specific directions for future research, perhaps the most intriguing of which would be a detailed investigation of the dynamics of noise determined variable coupling, which may have deep implications regarding the continuous transformation of one algorithm into another and may provide for the formulation of theoretically tractable definitions of such notions as approximately algorithmic behaviour.

THE NON-EQUILIBRIUM BOLTZMANN MACHINE

Many approaches have been taken to devising a theory of structures similar in their formal properties to nervous systems [Palm, 82], [Hopfield, 82], [an der Heiden, 80], [Caianiello, de Luca & Ricciardi, 67]. Most try to account for the behaviour of the networks by explicitly solving an enormous system of differential equations. While possible, at least in theory, this approach is reminiscent of an attempt to find the temperature and pressure in a room by explicitly solving the equations of motion of the molecules of air. Only recently has the theory of stochastic systems [Adomian, 83] received a wide following, so it comes as no surprise that most approaches are deterministic. An early exception [Sejnowski, 76] began investigating the expected behaviour of single neurons as opposed to the more usual analysis of the expected behaviour of the network as a whole. This work led directly to the Boltzmann machine.

A Boltzmann machine [Ackley, Hinton & Sejnowski, 85] is an abstract construct consisting of a set of nodes connected by links. At any point in time, the output of each node is either zero, in which case the node is said to be in the passive state, or one, in which case the node is said to be in the active state. In response to their input signals, the nodes fluctuate back and forth between these two states in a manner described below.

Nodes interact via connecting links. Each link has a

weight, which can take on any real value. The input of a node is the sum of the outputs of the nodes that drive it, each multiplied by the weight of the connecting link. No node drives itself, and the link weights are symmetric. These last two assumptions simplify analysis, but may be relaxed without damaging the formal properties.

The state of the system as a whole is given by the vector of node states, ζ_i, together with the matrix of link weights, W_{ij}. The energy of interaction between two nodes is taken to be W_{ij} if both are active, and vanishes if either node is passive. The total energy of the system is therefore given by

$$E = -\frac{1}{2}\sum_{i,j} W_{ij}\zeta_i\zeta_j \tag{1}$$

and the change in global energy due to a change of state of the i^{th} node is just

$$\Delta E_i = -\Delta\zeta_i\sum_j W_{ij}\zeta_j \tag{2}$$

A single parameter, T, controls the degree of determinism of the system, by virtue of its appearance in the decision rule giving the state of a node as a function of its input. This parameter plays the role of temperature and hence appears in the context of kT, where $k \approx 1.38\cdot10^{-23}$ J/K is Boltzmann's constant. The product kT gives, essentially, the amount of energy required to represent one bit of information at an absolute temperature of T. It is significant that throughout

all of theoretical physics, temperature appears only in this context.

Now, each node samples its input at a mean rate of once in every interval of length τ. Immediately after sampling, the node makes a decision as to which state it will enter. Independent of its current state, the node becomes active with probability

$$p_i = \frac{1}{1+\exp\left(\frac{\Delta E_i}{kT}\right)} \tag{3}$$

This is called a Boltzmann decision and the distribution of node states which results is none other than the Boltzmann distribution for a lattice gas. The mean firing rate of a node performing a sequence of such decisions is just

$$\omega_i = \frac{2\pi}{\tau} p_i \tag{4}$$

and encodes the information output of the node.

The sigmoid form of equation (3) and the nonlinear transfer function of a neuron are related, as are the rate code typically observed in nervous information transmission and that given by equation (4). Still, the expressions given above should be thought of as describing the expected operation of an ensemble of neurons. Hence a single node in a Boltzmann machine is an analog not of an isolated neuron, but of a neuron acted upon by a stochastic background of other neurons.

The Boltzmann machine performs its computations by the non-equilibrium process of settling to thermal equilibrium, which minimizes the Helmholtz free energy of the system, of which the energy in (1) is only one component. The other vital term is the entropy of the system, which measures the information content of the system, taking into account the energy of the mean node states and the energy contained in the fluctuations about the means. The significance of this fact can be seen by substituting the equivalent entropy,

$$\Delta S_i = -\frac{\Delta E_i}{kT} \tag{5}$$

into equation (3). The action of each node can thus be interpreted as a decision regarding the significance of the information, ΔS_i, appearing at its input, given the overall noise level of the information.

In settling to equilibrium, the system evolves towards one of its stable states; at equilibrium, the system will wander randomly about in state space near such a stable state. The size of the region in which the wandering takes place is determined by the temperature, T, shrinking to a single point at a temperature of zero. The number and location of the stable states is determined by the W_{ij}, and these points coincide with the local minima of the Helmholtz free energy,

$$F = E + kTS \tag{6}$$

the gradient of which provides the force driving the system

towards equilibrium. It is to be expected, therefore that the techniques of statistical mechanics provide a powerful formalism for analysis of the Boltzmann machine. In fact, the approach has been adapted from recent work on the properties of spin glasses, of which the Ising model of ferromagnetic domains is perhaps the best known example.

The entropy appearing in equation (6) is that of the system as a whole, and is therefore the expectation of the individual node entropies,

$$S = -\sum_i p_i \ln p_i \tag{7}$$

Now, the Boltzmann distribution is precisely that assignment of probabilities which remains maximally noncommittal about all other aspects of the system. It is found by minimizing equation (6) subject only to consistency conditions, such as requiring the probabilities to sum to unity. That is to say, the Boltzmann distribution is that deviation from equipartition of states that just encodes the information in the energy function of the system.

For this reason, it comes as no surprise that the nodes are performing statistical inference upon the information presented to them [Hinton & Sejnowski, 83]. It is fairly immediate that a Boltzmann decision is simply an application of Bayes' theorem. To see this, let a node be associated with a hypothesis, h, and let it be driven by a node that is associated with an item of evidence, e, that bears on the hypothesis. When the node is active, it is deemed to have

made the decision, h is true. Conversely, when the node is inactive, it is deemed to have made the decision, h is false. As the node fluctuates back and forth between these two states, it is continually revising its decision, based on the instantaneous value of its input signal. The probability that the node is active is thus an estimate of the probability that the hypothesis is true, based only on the information content of the evidence provided by the input signal of the node. Now, Bayes' theorem can be expressed in the form

$$p(h|e) = \frac{p(h)\,p(e|h)}{p(h)\,p(e|h) + p(\neg h)\,p(e|\neg h)} \tag{8}$$

That is,

$$p(h|e) = \frac{1}{1 + \dfrac{p(\neg h)\,p(e|\neg h)}{p(h)\,p(e|h)}} \tag{9}$$

but elementary probability theory provides a result that allows this expression to be rewritten in a most suggestive form. Namely, $p(b)p(a|b) = p(a \wedge b)$, so that

$$p(h|e) = \frac{1}{1 + \dfrac{p(e \wedge \neg h)}{p(e \wedge h)}} \tag{10}$$

A comparison with equation (3) leads one to identify $p(h|e)$ with p_i, and thus to write

$$\frac{p(e \wedge \neg h)}{p(e \wedge h)} = \exp\left(\frac{\Delta E_h}{kT}\right) \tag{11}$$

which equates a change in log probabilities with a change in

entropy by virtue of equation (5). Therefore, the increment of energy passed along from node e to node h, ΔE_h, causes it to adjust its firing probability by exactly right amount, if the probability of hypothesis h given evidence e is equated to the firing probability of node h. This same argument can be applied when e is a set of nodes so long as proper account is taken of any nonorthogonality in the output signals of these nodes.

Here, then, is the basis of inductive reasoning as well as deductive reasoning, due to the upper limit on the firing rates. A node is capable of representing absolute certainty at the extrema of its range, absolute uncertainty at the mid point, and any other degree of conviction elsewhere. Groups of nodes can reach a consensus by the process of settling to thermodynamic equilibrium, a process which requires only the most rudimentary and uniform behaviour on the part of each node, yet is capable of generating global behaviour of any required degree of complexity.

Modelling the Nervous System

Before a Boltzmann machine becomes a reasonable model of a nervous system, it must be able to interact with its environment. For this purpose, the nodes are grouped into three subsets. The first division is between that set of nodes, V, which are visible to the environment and able to interact directly with it, and the remainder, H, the set of nodes which are hidden from the environment and can only

interact with it via the nodes in V. The set V is further subdivided into sets I and O, the input and output nodes. The sets V and H are necessarily disjoint while I and O need not be.

In these terms, a single stage nervous system is one in which the set H is empty and in which I and O are identical. In a two stage nervous system , I and O each contain nodes not contained in the other, while H is still empty. A three stage nervous system has nodes in H as well as in I and O, which latter pair are not identical, though they may contain a common subset. This division of the nodes into subsets does not preclude connections between any pair of nodes, nor does it imply them. The three sets, I, H and O, are thus abstractions of the more usual division of neurons into sensory neurons, interneurons and motor neurons.

The environment directly drives the input nodes, delivering increments of energy to them via a wide variety of sensory modalities, and these energy increments are summed with any input from other nodes in accordance with equation (2). The nodes thus respond to the information content in the input from the environment in exactly the same way as to the information passed to them from other nodes. An important consequence of this fact is that the signals entering the system continually drive it away from equilibrium, while the system itself continually strives towards equilibrium.

The output nodes influence the environment by driving actuators such as muscles, glands and the like. In order to

simplify the present analysis, these actuators will be considered to be components of the environment. Thus the term environment does not refer to that of the organism, but to that of the nervous system. The organism is then a part of this environment.

Given any state of the environment discernible by the input nodes, there is an optimal configuration of the output nodes. That is, for any given stimulus, there is an optimal response, in the sense that this response leads to a maximal satisfaction of the organism's needs. In practice, the environment is itself a stochastic system, so that the optimal mapping between stimulus and response will be subject to random fluctuations about some fixed or relatively slowly varying operator.

Since the hidden nodes occupy a position between the input nodes and the output nodes, they determine the salient features of the mapping from input to output. It remains to be shown by what mechanism the system organizes itself such that the mapping converges to the optimal one, without requiring any guidance other than that which is implicit in the relationship between the output of the system and its subsequent input.

This question is taken up in appendix A, where the results are applied to the operation of the sensory preprocessing system of the organism. It can be seen from the discussion which appears there that learning of arbitrary high order constraints in the environment can be achieved in a

principled way by modifying local behaviour in response to global parameter, called effectiveness. The modifications are performed in a uniform way, completely independent of the problem domain, driven solely by the information which needs to be processed, without any *a priori* knowledge of what that information may be. The only limitations are imposed by the suite of sensory modalities and actuator capabilities. Thus a Boltzmann machine is an ideal model for a nervous system and provides precisely the sort of substrate for self organization that was sought in the introduction. Given a sufficiently large number of degrees of freedom, a Boltzmann machine can exhibit arbitrarily intelligent behaviour. It is premature to describe the behaviour of a small system, of the size of the nervous system of *C. elegans*, as intelligent. In this context, adjectives like adaptive and discerning are more appropriate even though the underlying principles appear to be the same.

Of greater interest in the context of the discussion of the motor postprocessing system is the way the dynamics of the Boltzmann machine is modified when the summing operator at the input of each node is replaced by a lossy integrator. The net results are that the signal to which the node responds is effectively a short time average of recent past inputs and that the response of the node is delayed by the time constant of the integrator. Nontrivial time dependent behaviour is introduced by this mechanism and by the relaxation of the constraint that the link weights be symmetric. It also

permits the analysis of non-equilibrium behaviour, in that the temperature may vary from node to node in a manner determined solely by the statistical properties of the signals.

The argument which follows is a brief sketch of the line of reasoning developed in detail in appendix B. With this sketch in hand, it is possible to proceed directly to the discussion of the motor postprocessing system. However, the reader who is unfamiliar with the type of arguments outlined in the remainder of this section may wish to review appendix B before proceeding to consider their application to the analysis of the motor postprocessing system.

Equation (2) expresses the change in the energy of the system associated with a change of state of the i^{th} node. The sum on the right hand side of that equation is therefore a force. In order to address the motion of the system through state space in response to this force, the time constant τ, introduced by the membrane resistance and the membrane capacitance of the neurons which the nodes model must be taken into account. The time dependent form of equation (2) is thus

$$\tau \frac{d\eta_i}{dt} + \eta_i(t) = f_i(t) \qquad (12)$$

where η_i is the effective membrane potential of the i^{th} node and f_i is the force exerted by its input signal. Another way to interpret this expression is that f_i is the instantaneous membrane potential while η_i is a moving average of f_i which is exponentially smoothed using a time constant of τ. The

solution of this equation is found to be

$$\eta_i(t) = \eta_i(0)e^{-\frac{t}{\tau}} + \frac{1}{\tau}\int_0^t e^{\frac{s-t}{\tau}} f_i(s)\,ds \tag{13}$$

by the method described in appendix B.

Up to this point, ζ_i has represented the instantaneous output value of a node, which can take on the values 0 and 1, and which takes on the latter value with probability given by equation (3). In order to effect the transition to a continuous, time dependent version of the model, it is useful to let this symbol henceforth denote the effective value of the stochastic binary variable. That is, over an interval on the order of τ, the net effect of ζ_i on the motion of the system can be found by treating it as though it were a continuous random variable whose mean is given by the right hand side of equation (3) and whose variance may easily be verified to be $p_i(1 - p_i)$. Since its statistics are non-stationary in the time dependent model, the mean and variance of the output of each node are functions of time.

Now, letting ξ_i denote the effective value of the input signal to the i^{th} node, defined analogously to the new interpretation of ζ_i but originating outside of the system, the force acting on each node is just

$$f_i = \xi_i + \sum_j W_{ij}\zeta_j \tag{14}$$

which, being a weighted sum of random variables, is itself a

random variable.

The random fluctuations superimposed on f_i are, in general, different for each node. Therefore the single global temperature used in the model presented above must be replaced with a vector of equivalent noise temperatures, which are shown in appendix B to be proportional to the root mean square value of the fluctuations, denoted by σ_i. By defining the temperature of each node in terms of the statistics of its input signal, the approach to equilibrium of the system as a whole becomes determined by its equation of motion and there is no need to introduce such artificial devices as a cooling schedule.

It is argued in appendix B that the time behaviour of the system is most easily approached by considering the variations of the signals about their equilibrium values. In so doing, one finds that the effective coupling between the nodes differs from that given by the link weights, due to the fact that the slope of the sigmoid transfer function of each node is a function of the equivalent noise temperature of its input signal. The effective transfer operator of the system, A, is thus given by

$$A_{ij} = \delta_{ij} - \frac{W_{ij}}{\sigma_j} p_j(1-p_j) \qquad (15)$$

where the second term on the right hand side is the interaction matrix, as derived in appendix B. This is an important step, because it allows stochastic signals in a

fixed structure nonlinear system to be treated like deterministic signals in a variable structure linear system.

With these ideas in hand, the linearized equation of motion of the system is just

$$\tau \frac{d}{dt}(\eta - \langle\eta\rangle) + A(\eta - \langle\eta\rangle) = \xi - \langle\xi\rangle \tag{16}$$

which has the solution

$$\eta(t) - \langle\eta\rangle = T(t)(\eta(0) - \langle\eta\rangle) + \int_0^t T(t-s)(\xi(s) - \langle\xi\rangle)\,ds \tag{17}$$

where the matrix T(t) satisfies a matrix differential equation given in appendix B.

In analyzing the time dependent behaviour of a Boltzmann machine subject to these generalizations, the usual definition of statistical expectation applies only to the equilibrium values of the signals. Thus, for example, $\langle\eta\rangle$ is the equilibrium value and η is the effective value of the membrane potential.

In appendix B, all these ideas are developed in detail and the results applied to the central nervous system of the organism. The main idea that emerges is that the behaviour of the system can be treated as linear in certain parts of its state space, while transitions between one effectively linear regime and another occur in other parts of state space. In essence, a number of linear systems are embedded in the nonlinear one. Changes in the input can thus drive the system

from one dynamical regime into another, where each regime obeys qualitatively different dynamics from all the others. This is based on the mechanism of noise determined variable coupling which is related, in appendix B, to the idea of context sensitive association.

The same idea is used in the analysis of the motor postprocessing system to explain how the central nervous system brings about a transition between two different dynamical regimes of the motor postprocessing system. In that case, one effective structure damps out fluctuations in its input, while the other amplifies them until it achieves sustained oscillations which drive the somatic musculature and produce the serpentine motion typical of *C. elegans*.

THE STRUCTURE AND PHYSIOLOGY OF *C. ELEGANS*

C. elegans is a free living soil nematode, approximately 60 microns in diameter and almost a millimetre in length. Figure 2.1 shows the organism in cross section. The exterior of the organism is coated with a cuticle that is excreted by the hypodermal cells. It forms the base for the attachment of muscle, which is distributed rather than localized at specific points of attachment. The cuticle is punctured anteriorly by the trilaterally symmetric mouth and ventro-posteriorly by the anus. In the hermaphrodite, the vulva is located ventrally about 60 percent of the way from the mouth to the tip of the tail. Approximately one in a thousand larvae become males, with a copulatory spicule replacing the vulva, although it is located much closer to the anus. The only other major opening in the cuticle is the excretory pore, located ventrally just posterior to the head. In addition, of course, there are several sensilla that have microscopic openings providing access to the external chemical milieu.

The mouth opens into the buccal cavity, which forms the anterior part of a double bulbed pharynx. The second bulb houses a three toothed grinding and filtering organ, through which food must pass before entering the intestine via the pharyngeal-intestinal valve. The pharynx is contained in a basement membrane which separates it from the rest of the organism. Only two neurons pass through this membrane and connect to the animal's main nervous system, although there

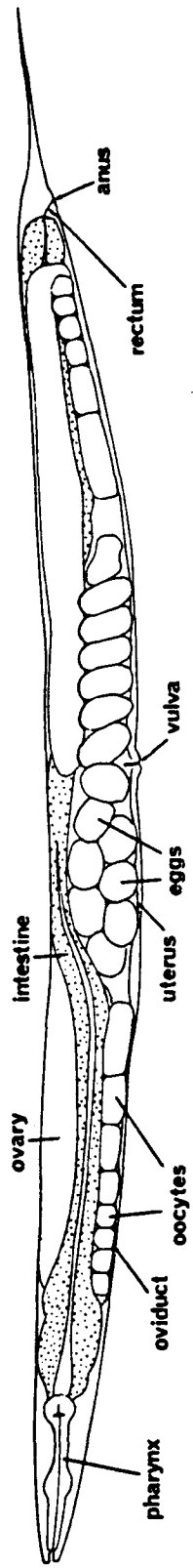

Figure 2.1 - The main anatomical features of *Caenorhabditis elegans*

(Adapted from Wood, [Chalfie & White, 88])

are a total of 20 neurons innervating the pharynx, which forms an autonomous subsystem [Albertson & Thomson, 76].

The muscles are located in two pairs of bundles closely apposed to the hypodermis. The dorsolateral and ventrolateral pairs are innervated identically, so that the organism is mobile in the dorsoventral plane only. The head, however, has an additional degree of freedom with the ability to move from side to side. The pseudocoelom, a large annular cavity between the intestinal tract and the muscles, is maintained at a turgor pressure of a few psi above ambient. This forms a hydrostatic skeleton [Lee, 65] against which the muscles act.

Movement is caused by dorsoventral undulations that propagate along the length the organism, from head to tail or vice versa, depending on the direction of motion. Such travelling waves are only possible if the contraction of one pair of muscle bundles is accompanied by lengthening of the opposing pair. The force causing this stretching is due to the turgor pressure. The cuticle is composed of several layers, the most important of which is a matrix of collagen fibres crossing at about 135 degrees [Crofton, 66]. An increase in the turgor pressure causes the enclosed parallelograms become flattened. As a consequence, the diameter of the worm is rather independent of the pressure, while the length is strongly dependent on it. It is by virtue of this structure that the muscular contractions lead to flexion instead of distension of the cuticle.

This method of locomotion is uniquely suited to the

physical environment in which *C. elegans* exists. Living in thin films of water in the interstitial spaces between soil particles, the organism is held against one of them by surface tension. The film forms a meniscus which meets the body of the animal at an angle determined by the depth of the film. When the body is displaced laterally by muscular contractions, the film is distorted and acts to oppose further displacement. This, together with friction between the cuticle and the substratum provides the reaction to the force exerted by the worm [Croll, 70]. Consequently, when the film is much thinner than the diameter of the animal, movement is accomplished with great difficulty due to the large normal force, and when the film is much thicker than the diameter, movement is difficult due to the lack of a restoring force. In the intermediate range, the rate of locomotion is determined by the difference between the rate at which the somatic waves move along the body of the worm and the slip between the cuticle and the substratum.

Various sensory modalities are available to *C. elegans*. Proprioreceptive input allows the nervous system to monitor such internal variables as hunger, fatigue, age, pressure, and posture. The latter two are due to stretch receptors distributed throughout the animal, while the others may be mediated by neurohumoural factors which are at present not well understood. It is possible, however, that feeding is controlled by stretch receptors, because expansion of the intestine due to intake of food results in lengthening of the

organism and defecation causes shortening. Also, since the nervous system continues to develop as the organism matures, moulting and the onset of reproductive behaviour could be controlled by the appearance of new structures and connections in the nervous system rather than receptors sensing changes in endocrine balance. Hence neurohumoural proprioreception may play a relatively minor part in the determination of behaviour. Nonetheless, it may perform a regulatory function such as controlling the overall level of arousal and rate of adaptation of the nervous system.

With respect to external variables, *C. elegans* exhibits chemoreceptors, thermoreceptors and mechanoreceptors. Also, an orthokinetic response to light has been reported, though *C. elegans* does not possess obvious photoreceptors. Most of the sensilla are located in the head. Six inner and six outer labial sensilla provide chemo- and mechano-sensory input from the region surrounding the mouth. Four cephalic sensilla are most likely mechanoreceptors analogous to the cephalic papillae of other nematodes. A pair of amphids are located just anterior to the nerve ring and are probably responsible for the chemo- and thermo-taxes. Because of the small spacing between them and the sensitivity of the organism to concentration and temperature gradients, it is thought that these sensilla operate klinotactically rather than tropotactically [Ward, 76]. Two deirids are located in the cervical region, and two more are found in the tail. A number of neurons in the head are not associated with any sensilla

and must therefore be proprioreceptors. The animal is sensitive to touch along its entire length, although no specifically sensory neurons are located near the cuticle. Somatosensory input is thus most likely mediated by the motoneurons in this region, which also act as stretch receptors. Finally, two phasmids are located just posterior to the anus, and are similar to the amphids, although less prominent. The outer labial and cephalic sensilla closely resemble insectine campaniform sensilla [Mill, 82].

The nervous system consists of 302 neurons in the adult hermaphrodite, which is the dominant free living form. The male has additional neural tissue, most of it located in the tail. In the first stage larva, L1, 212 of these neurons are present. Those which develop later are associated primarily with reproduction and mating. In this paper, only that part of the nervous system present at the L1 stage are taken into consideration.

The majority of the central nervous system is located in the circumpharyngeal nerve ring, which nestles between the two bulbs of the pharynx. Most of the anterior sensilla send process tracts rearward into the neuropil of the ring, although the amphidial tracts continue past it, and enter the ventral ganglion via the amphidial commissures. This ganglion is located immediately posterior to the nerve ring and, together with the retrovesicular ganglion, forms the anterior terminus of the ventral nerve cord. The ventral cord projects along the length of the animal towards the tail, periodically

sending commissures around the pseudocoelom into the dorsal cord, which terminates anteriorly in the nerve ring. The ventral cord terminates posteriorly in the pre-anal ganglion, while the dorsal cord projects into the tail spike and sends processes into the pre-anal ganglion via the lumbar commissures. These connect the pre-anal and lumbar ganglia, which receive input from the posterior deirids and the phasmids. The pre-anal ganglion also controls defecation. Motor output from the nerve ring controls pharyngeal pumping and the muscles of the head. The somatic musculature is controlled exclusively by the dorsal and ventral cords, and is instrumental in propagating waves down the body of the animal. The muscle cells are peculiar in that they send processes into the nerve cords instead of receiving processes from them, as is usual in most organisms.

Motor Postprocessing: The Noise Driven Filter

The motor postprocessing system is organized into two nerve cords: one located dorsally, innervating the dorsolateral muscle bundles, and one located ventrally, innervating the ventrolateral muscle bundles. Each cord also contains the processes of sensory neurons and neuronal processes coupling the nerve ring with the pre-anal and lumbar ganglia. With the exception of the distal stretch receptive endings of such motor neurons that possess them, these processes are irrelevant to the present analysis.

There are 58 motor neurons distributed between the two cords. These fall into six classes: DA, DB, DD, VA, VB and VD. While all of the cell bodies are located in the ventral cord, the classes DA, DB and DD send processes into the dorsal cord via a number of commissures distributed along the length of the organism. The dorsal muscles are exclusively innervated by DA, DB and DD while the ventral muscles are exclusively innervated by VA, VB and VD. The three pairs of classes, (DA, VA), (DB, VD), and (DD, VD) are functionally homologous.

The class DA has 9 members, labelled DA1 to DA9 from anterior to posterior, while VA has 12 members, VA1 to VA12. These neurons have distal processes projecting anteriorly beyond the region in which synapses, neuromuscular junctions and gap junctions occur, and it is assumed that these serve as stretch receptors. Without this assumption, the model discussed in this section does not work. The DA and VA motor

neurons are necessary for normal reverse locomotion, that is, for forward propagation of a wave of flexion along the body of the worm. These neurons form excitatory synapses and neuromuscular junctions mediated by acetylcholine. They receive excitation from the nerve ring primarily via the interneurons of classes AVA, AVD and AVE.

The classes DB1 to DB7 and VB1 to VB11 have distal processes projecting anteriorly beyond the regions in which connections occur, and it is again necessary to assume that these act as stretch receptors. These neurons are also cholinergic and are necessary for normal forward locomotion, or reverse wave propagation. Their excitation comes from the ring interneurons of classes AVB and PVC.

In the classes DD1 to DD6 and VD1 to VD13, the regions in which neuromuscular junctions occur are dorsoventrally opposed from the regions having synapses and gap junctions. DD are postsynaptic to VA and VB in the ventral cord and innervate the dorsal muscles while VD are postsynaptic to DA and DB in the dorsal cord and innervate the ventral muscles. The neuromuscular junctions are mediated by the inhibitory neurotransmitter, gama-aminobutyric acid, and these classes act as dorsoventral cross inhibitors. They are required for normal motion in either direction.

Of the four muscle bundles, three contain 24 muscle cells and the fourth contains 23. Within each bundle, the muscle cells are grouped into 12 pairs, with the cells in each pair largely overlapping. The first two pairs in each bundle are

located in the cervical region of the animal and are innervated by the ring motor neurons. The somatic musculature can thus be divided into ten more or less distinct sections of four dorsal and four ventral muscles each.

The spatial organization of the motor postprocessing system and somatic musculature is shown in figure 5.1. Due to the unavailability of complete and accurate biological data, this figure is already somewhat idealized in that it assumes that the members of each motor neuron class are distributed evenly along the length of the organism. Although this is not, in fact, the case, the model developed below seems to work quite well nonetheless. As well, the slight asymmetry introduced by the "missing" muscle cell is neglected, again without doing serious violence to the model.

The interconnections among the motor neurons and the muscles are shown in figure 5.2. Figure 5.2 a shows all the connections. Since the nerve cords were not completely reconstructed by the researchers who provided the source data for this analysis [White, et al, 1986], the structure had to be extrapolated into the unreconstructed regions for purposes of analysis. The connections actually reported in White, et al, are shown in figure 5.2 b and those which were extrapolated are shown in figure 5.2 c. Figures 5.3 a, b and c respectively show the connections within the cords, between the cords and to the muscles. Figures 5.4 a through f show the connections specific to each of the six classes of motor neurons.

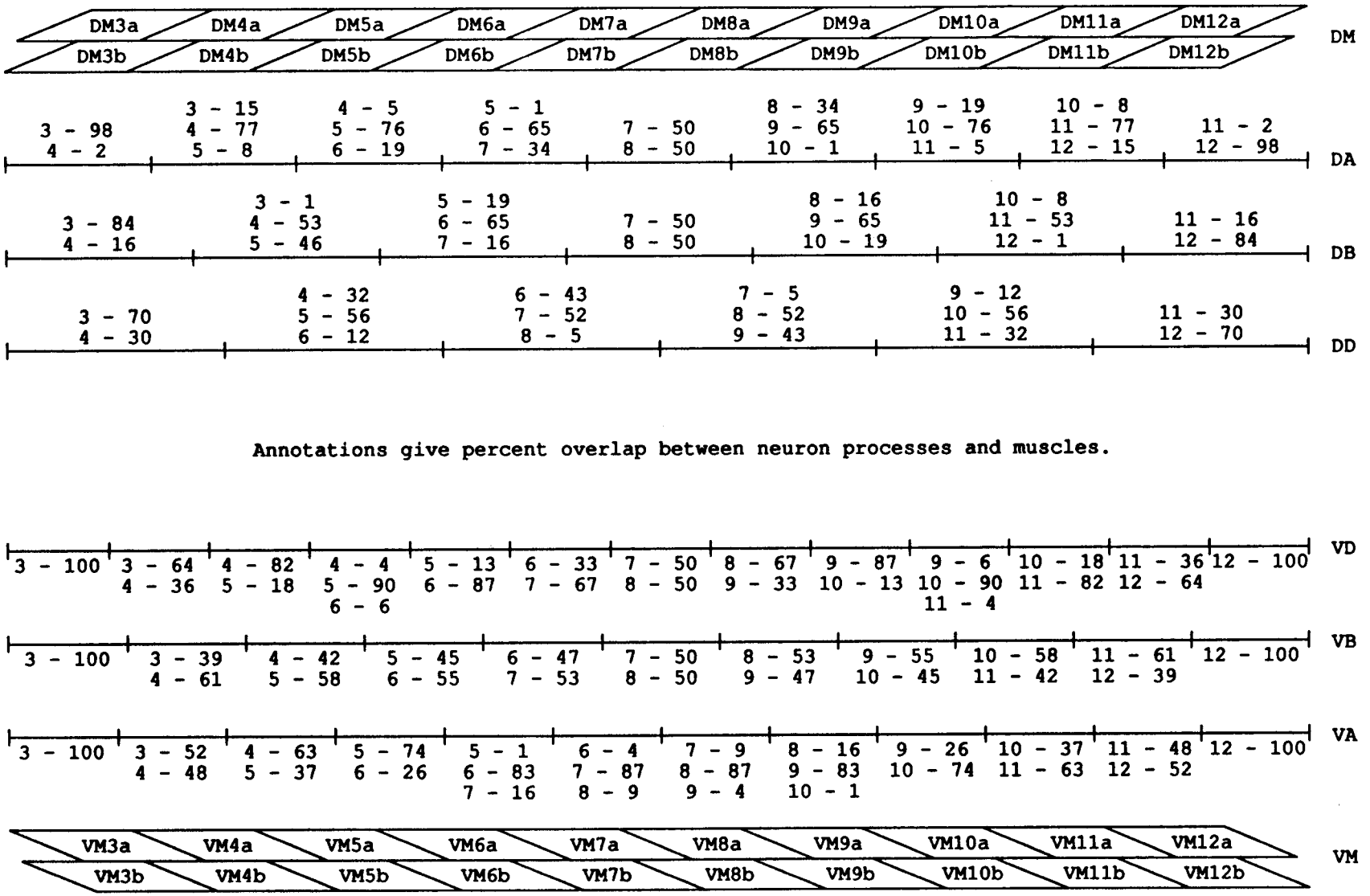

Figure 5.1 - Spatial organisation of the motor nervous system and somatic musculature

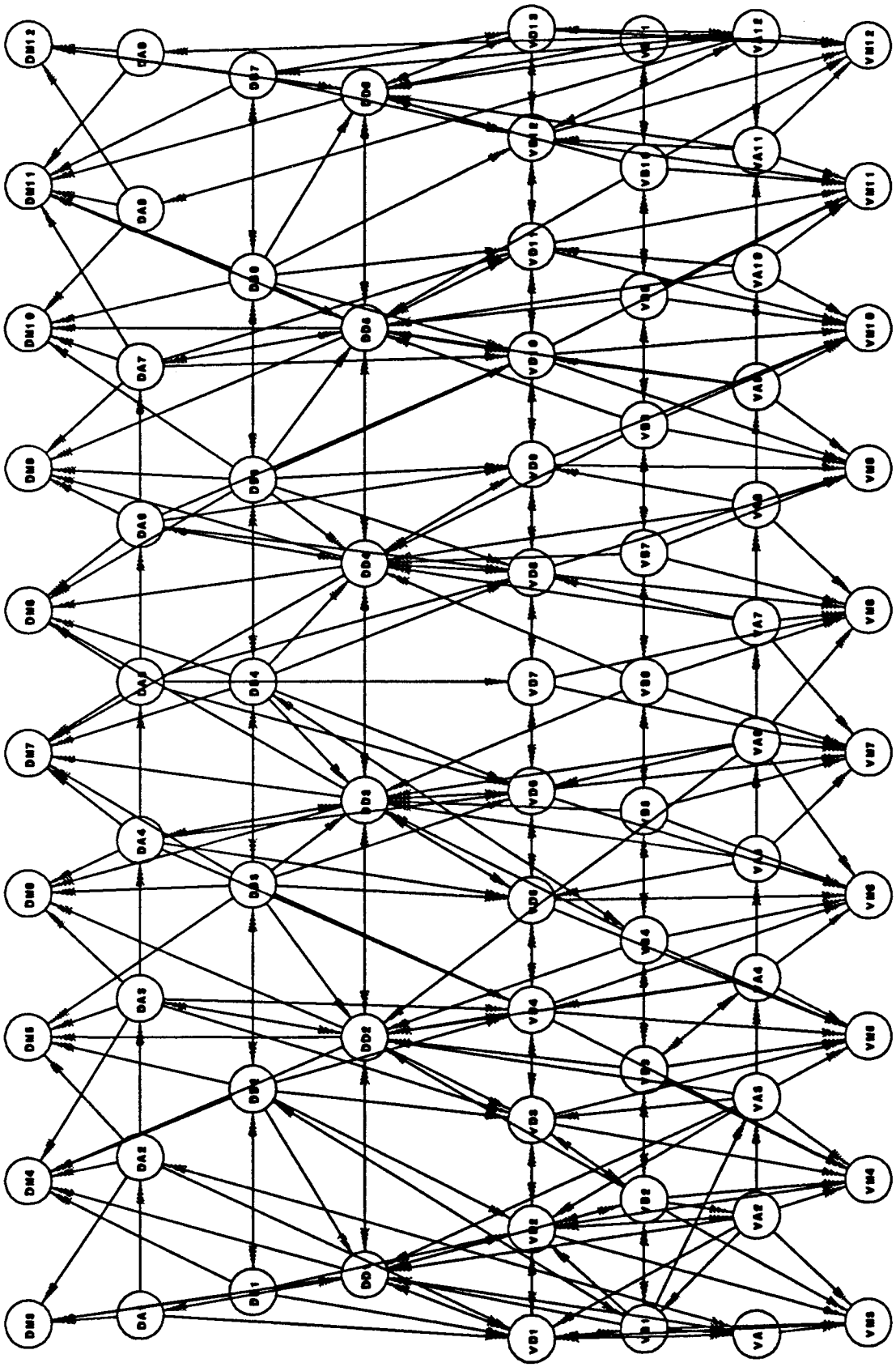

Figure 5.2 a - Connectivity of the motor nervous system and somatic musculature

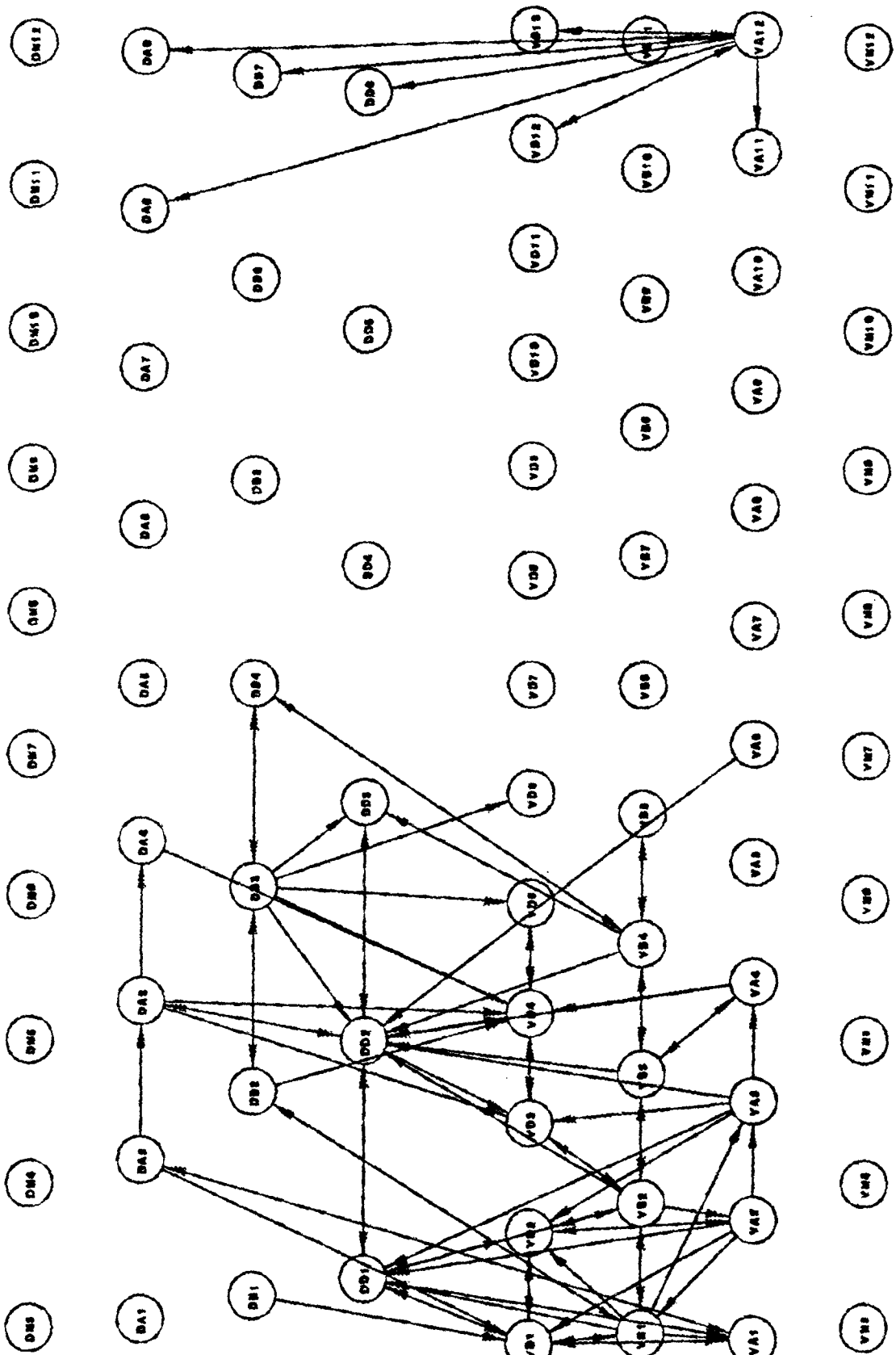

Figure 5.2 b - Connections reported by White, et al

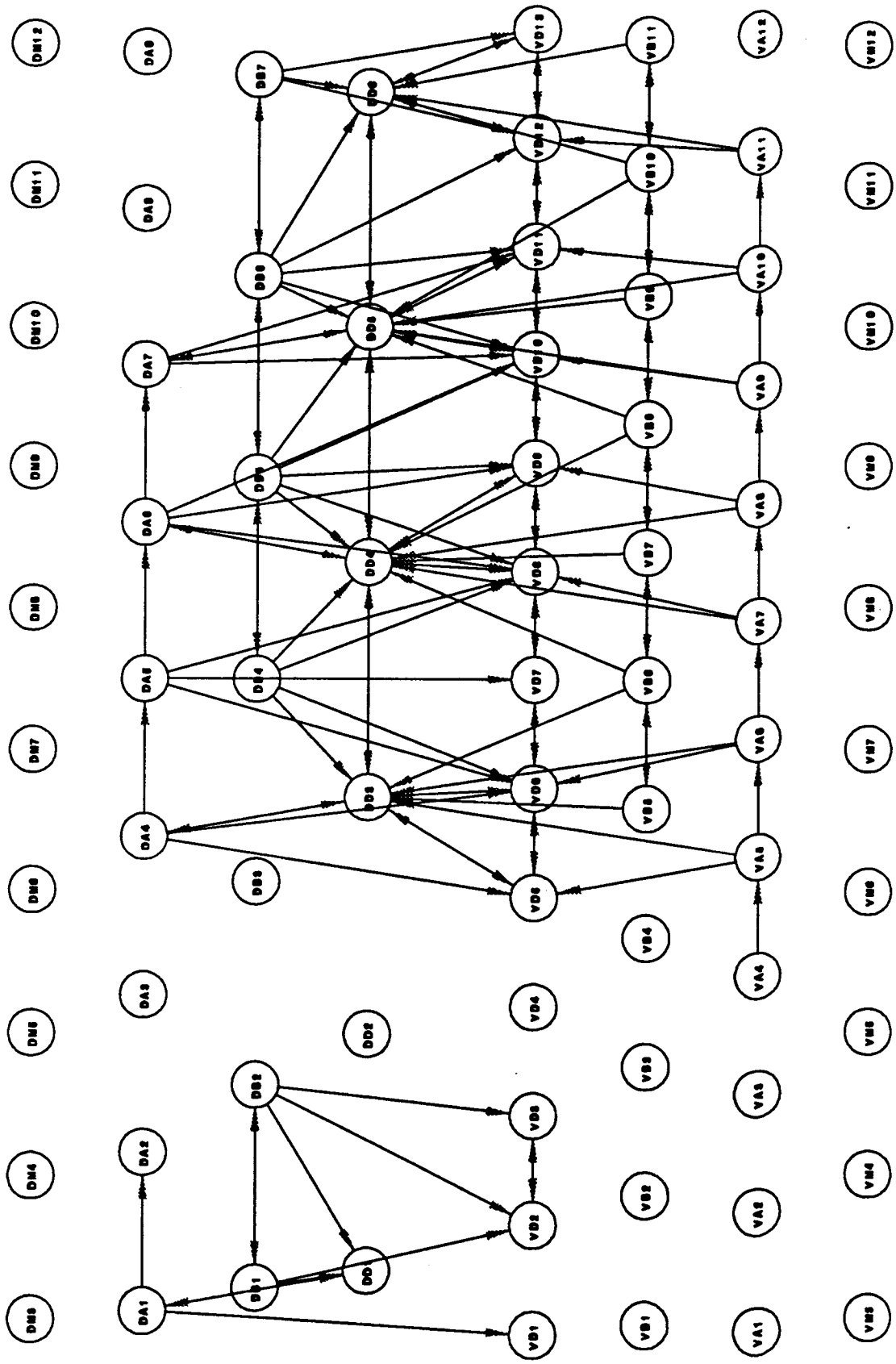

Figure 5.2 c - Extrapolated connections

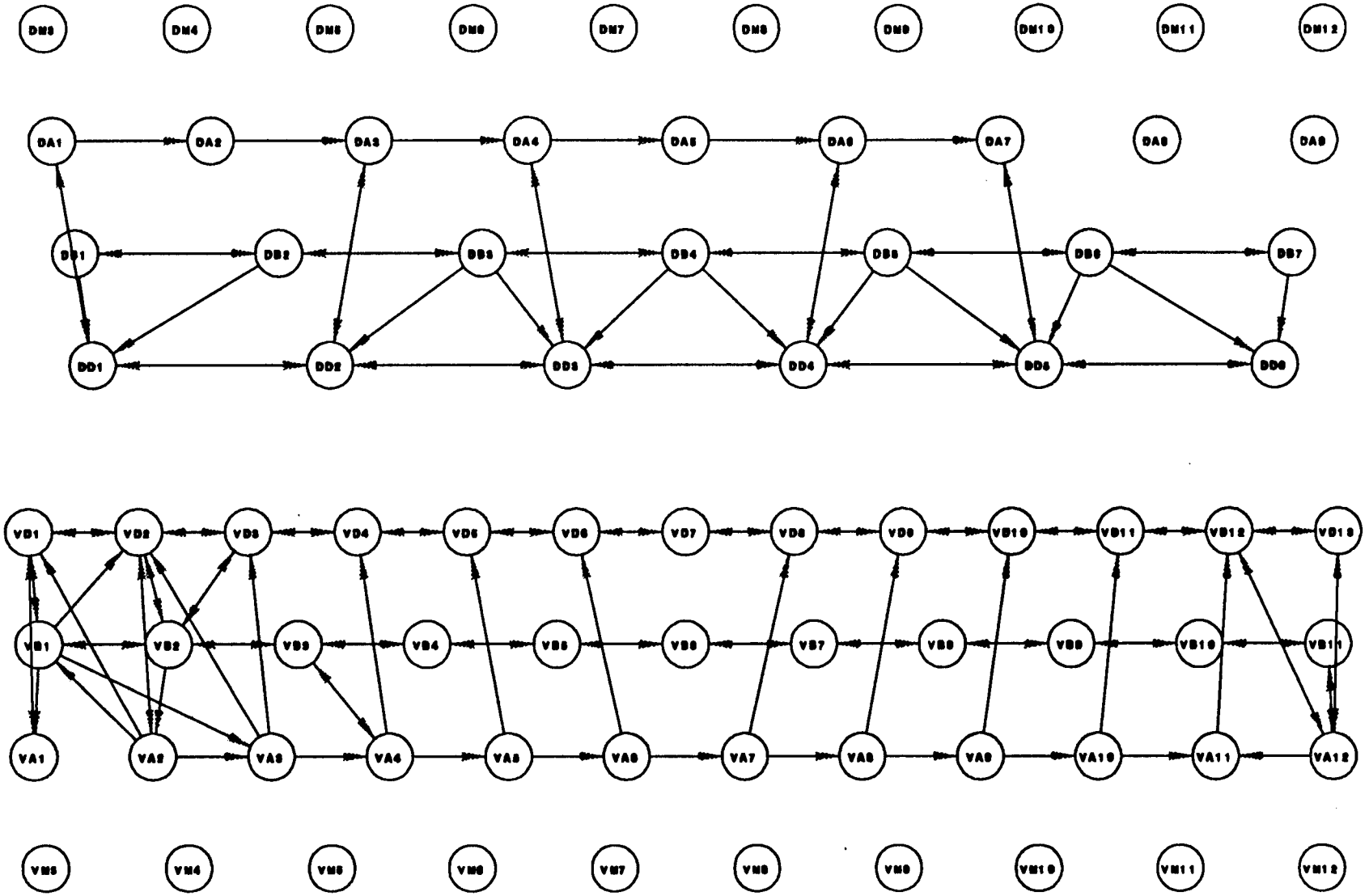

Figure 5.3 a - Connections within the dorsal and ventral nerve cords

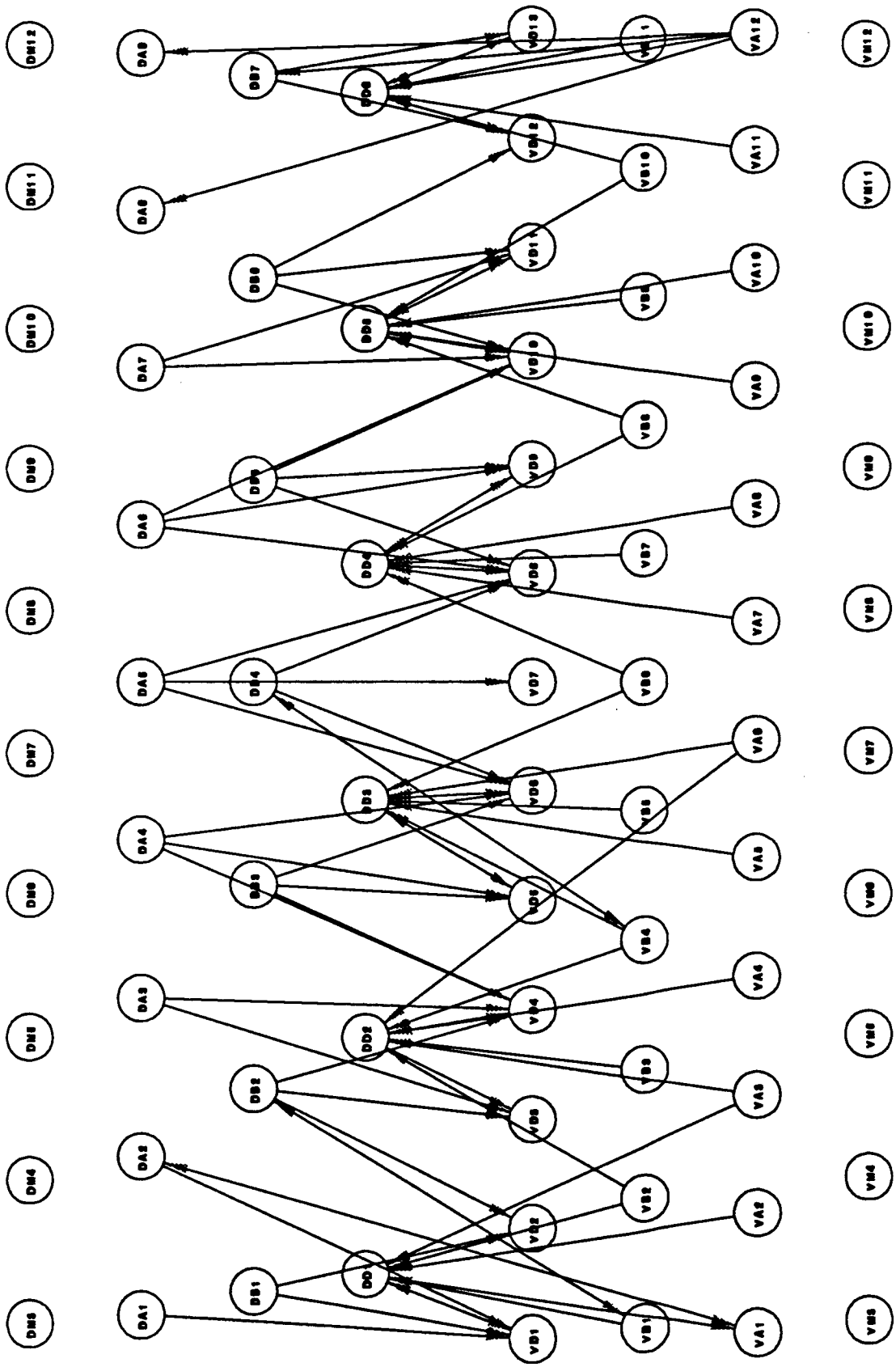

Figure 5.3 b - Connections between the dorsal and ventral nerve cords

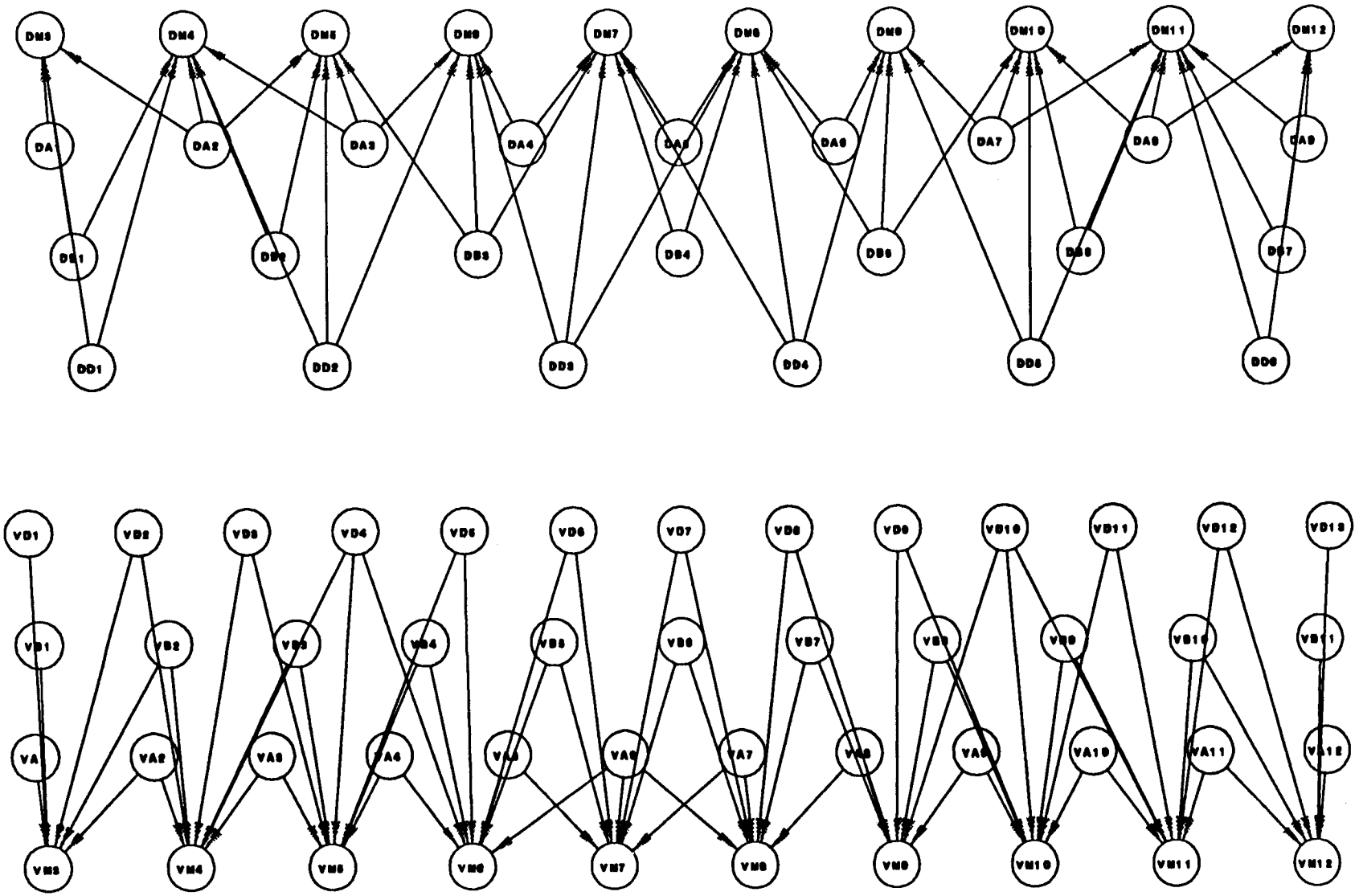

Figure 5.3 c - Neuromuscular junctions (extrapolated)

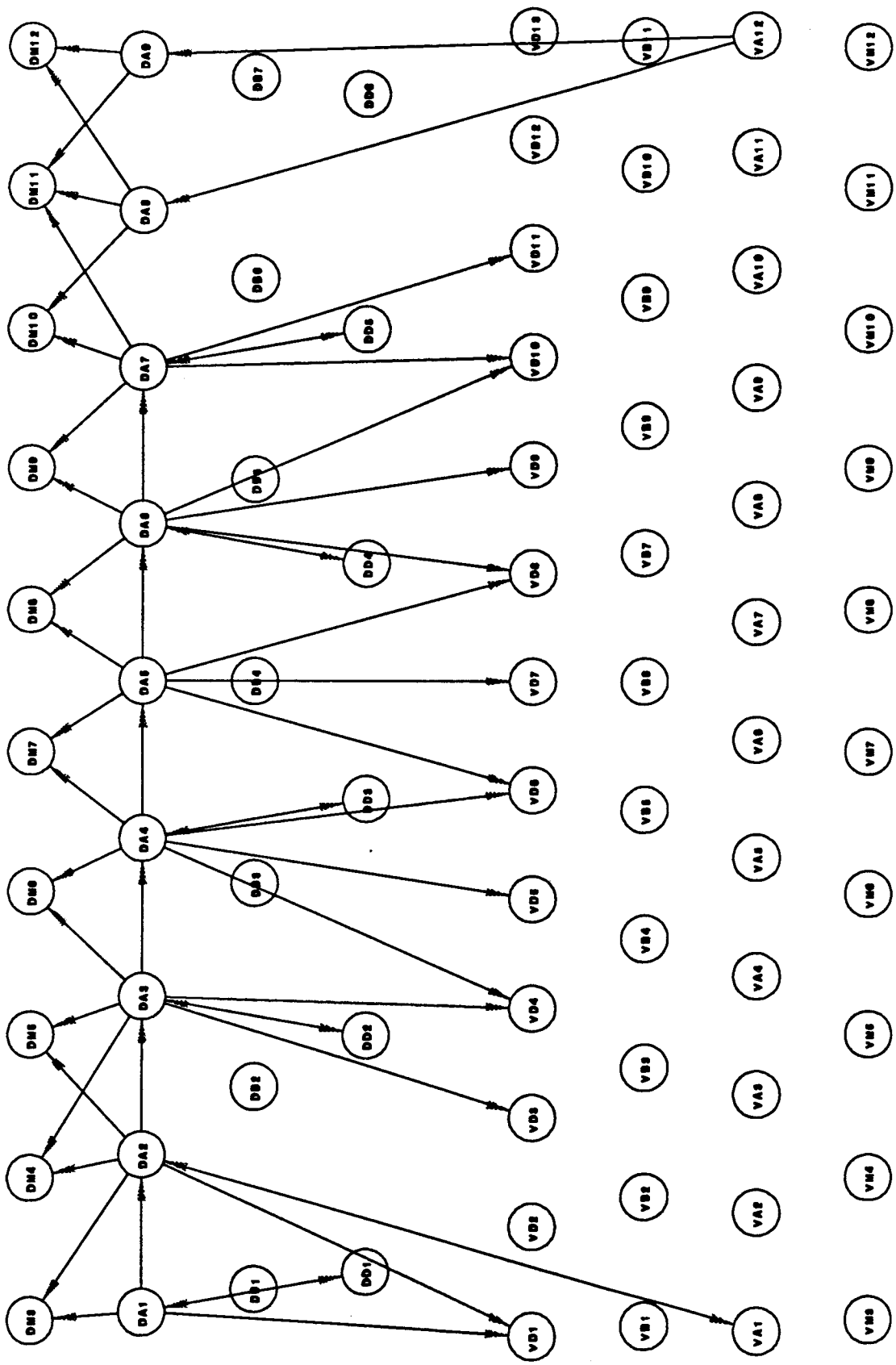

Figure 5.4 a - Connections of the DA motor neurons

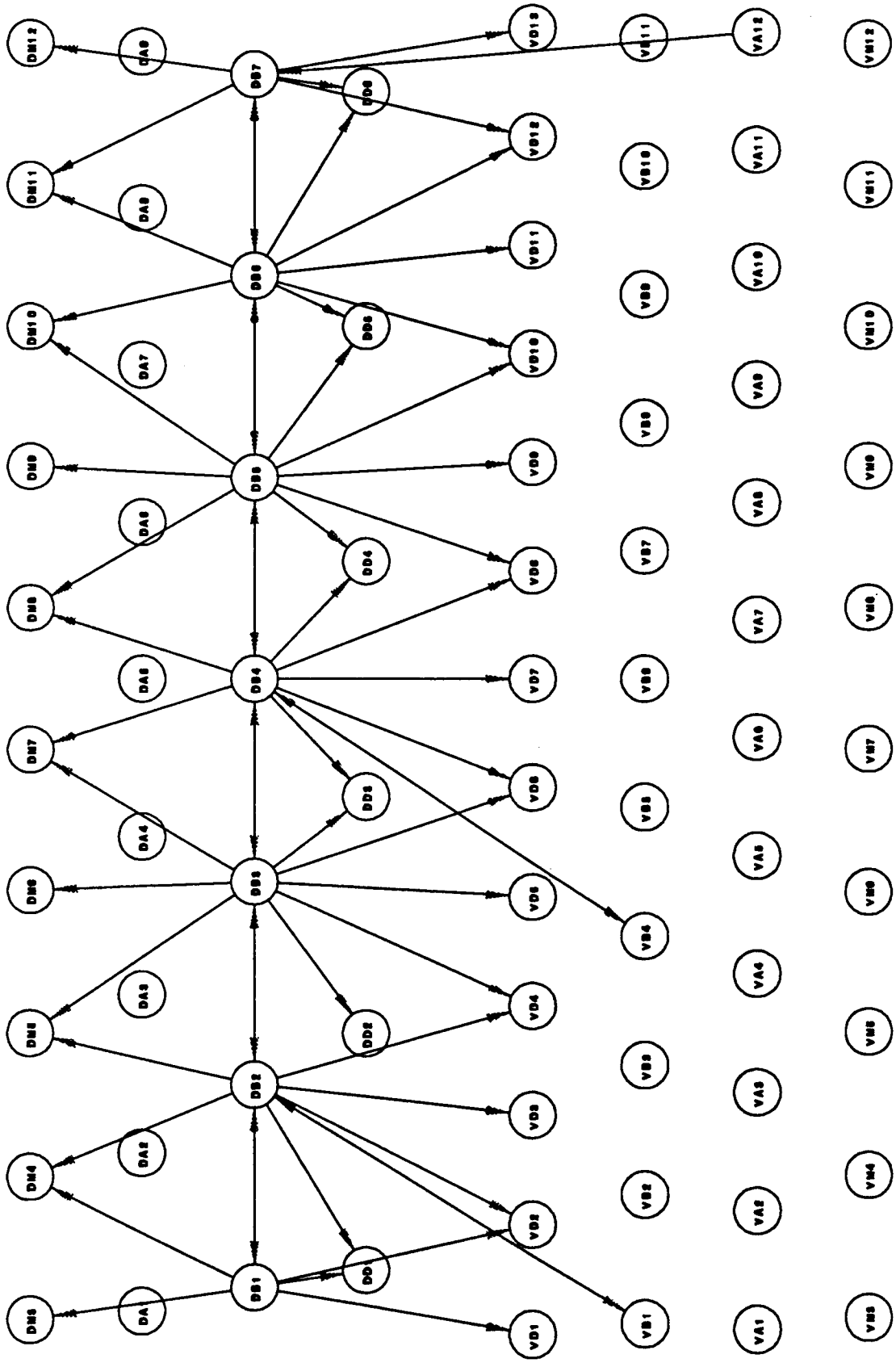

Figure 5.4 b - Connections of the DB motor neurons

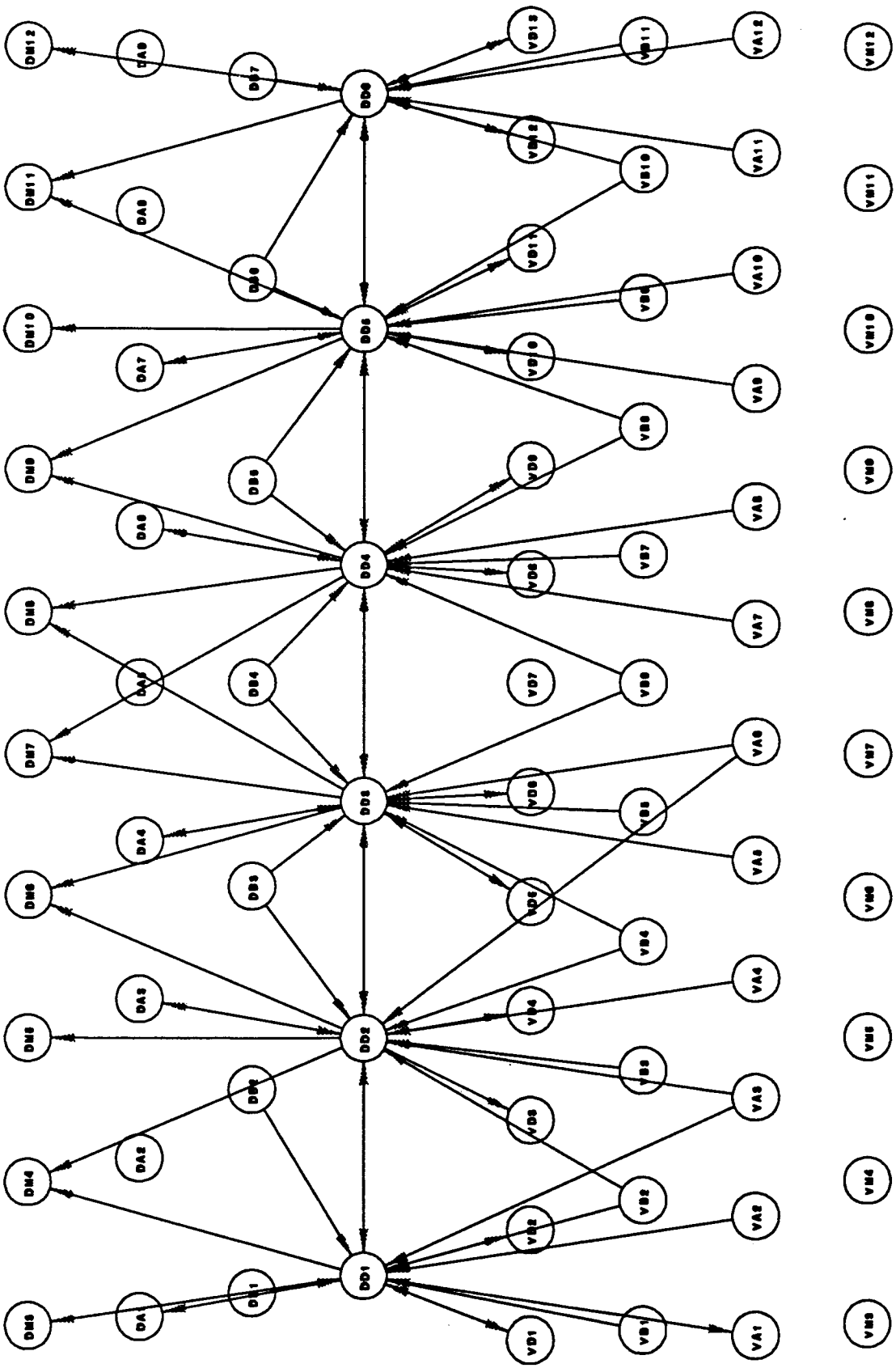

Figure 5.4 c - Connections of the DD motor neurons

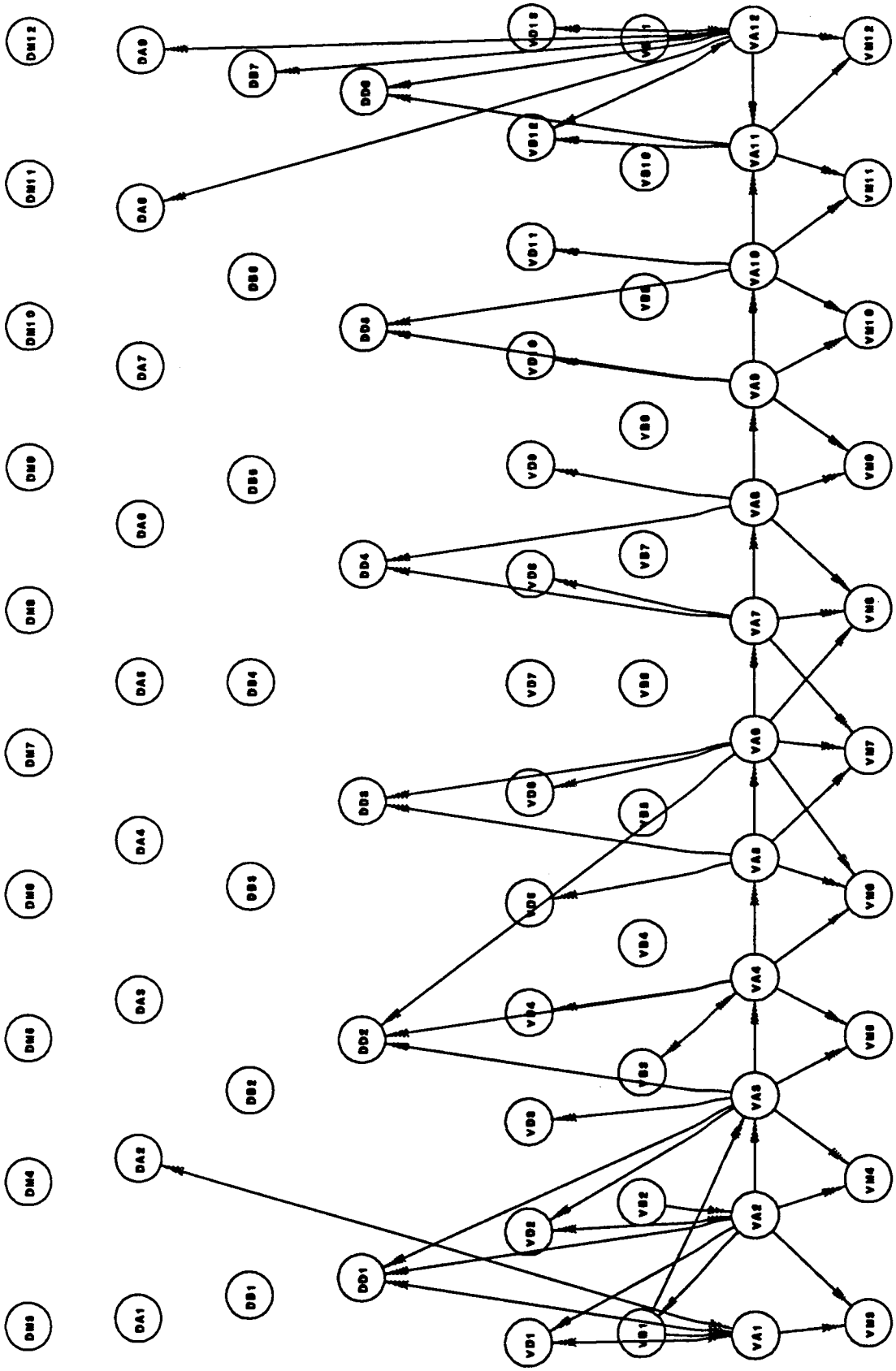

Figure 5.4 d - Connections of the VA motor neurons

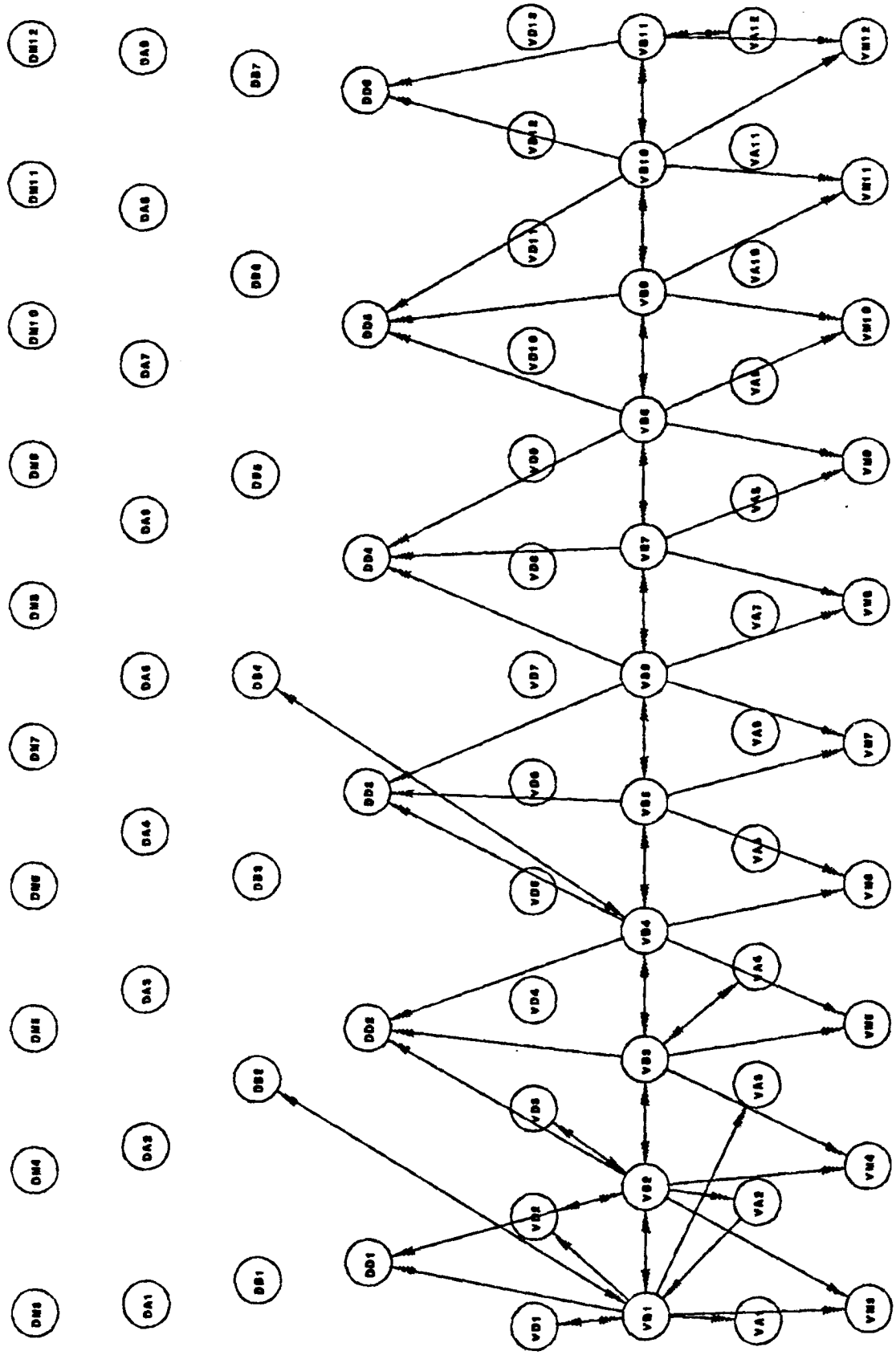

Figure 5.4 e - Connections of the VB motor neurons

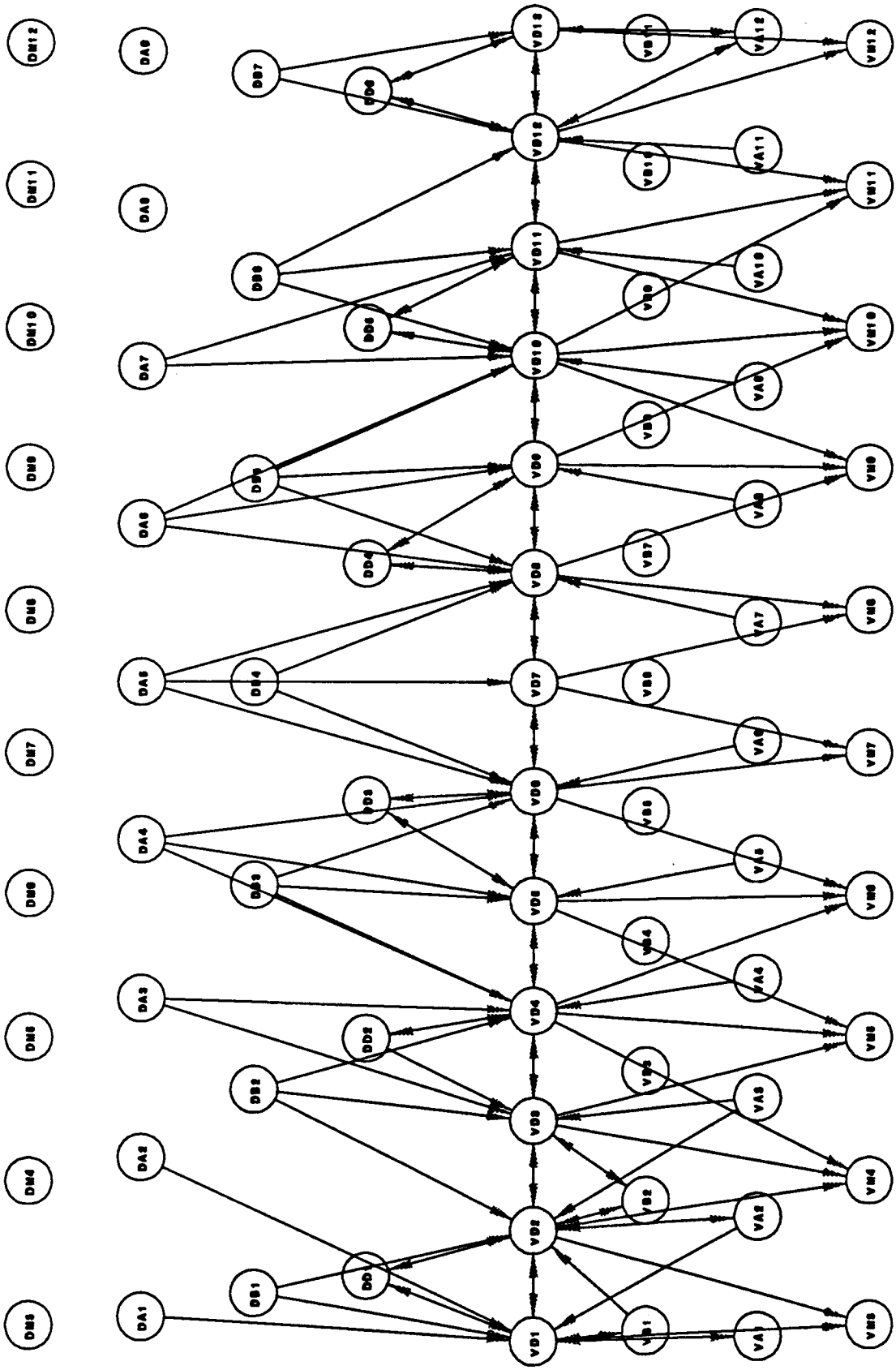

Figure 5.4 f - Connections of the VD motor neurons

The mechanical and topological structures depicted in these figures are quite complex and rendered a detailed analysis intractable without making recourse to some simplifying assumptions. The most important of these is that the dynamics of the motor postprocessing system can be approximated by those of an idealized system of ten sections, each having the structure shown in figure 5.5.

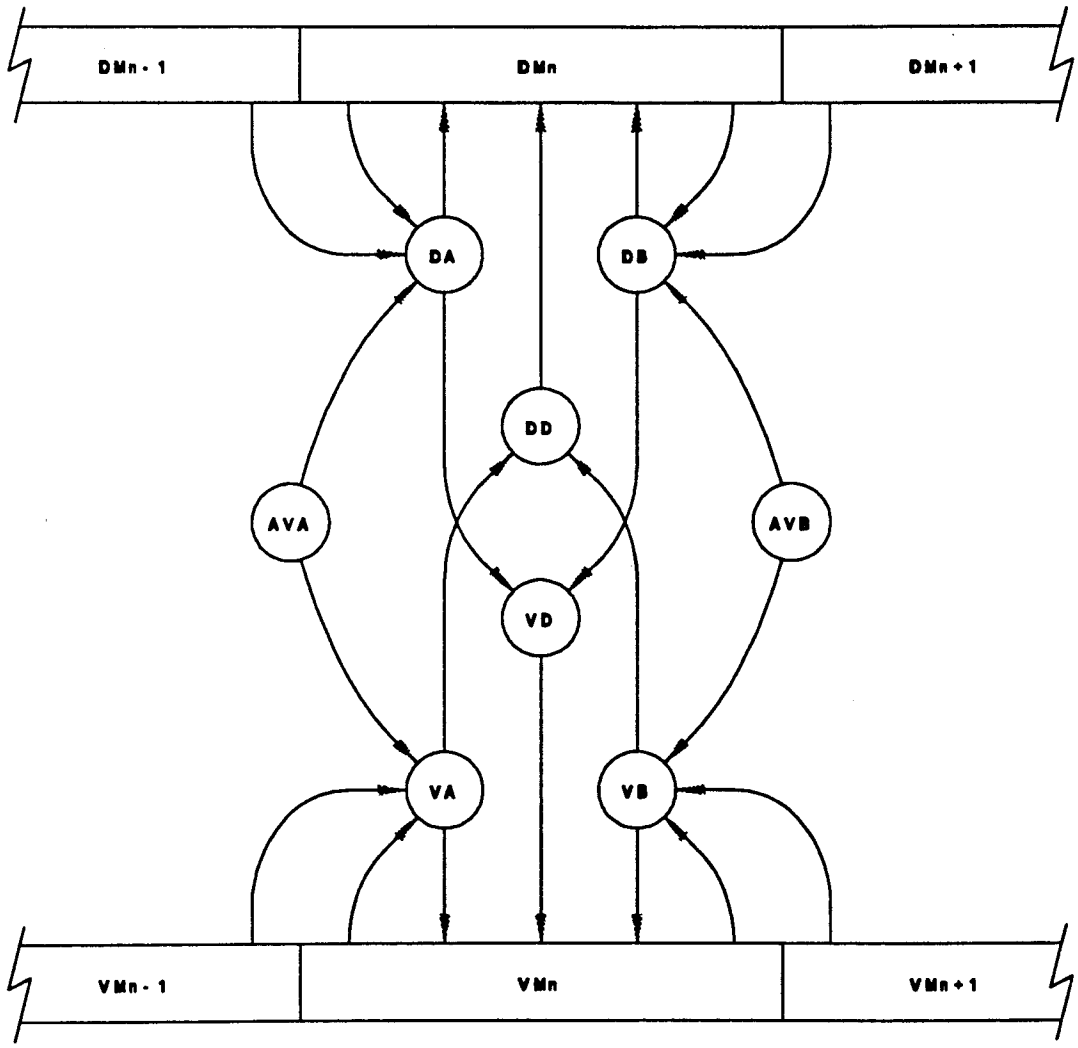

Figure 5.5 - Idealized structure of one section

This system can be thought of as two pairs of coupled oscillators, with only one pair active at any time, as determined by excitation from the nerve ring. At any given time, the same pair is active in all sections. One pair is composed of an oscillator made up of DA and DM and another oscillator made up of VA and VM, with coupling provided by DD, VD and the mechanical interactions between DM and VM. The other pair has a homologous structure incorporating DB and VB in place of DA and VA. The pair incorporating DA and VA is coupled to the previous section via the distal stretch receptive endings of these neurons, while the pair incorporating DB and VB is coupled to the next section. This coupling imposes a phase shift between adjacent sections resulting in the propagation of waves of dorsoventral flexion either forward or backward along the length of the body, depending on which pair is activated.

In this structure the only source of coupling between the sections is via the distal processes of DA, DB, VA and VB. The assumption that the dynamics of this system are approximately equivalent to those of the real system rests on the idea that the interpenetration of sections in the real system complicates the coupling in such a way that the mean phase shift between sections is relatively similar in the real and idealized systems.

To see how this system functions, consider what happens when DB is excited by the nerve ring. Its output increases and provides drive to the dorsal muscle cells of the section.

After some time delay, the muscle begins to contract in response to the drive. The contraction causes a decrease in the signal provided by the stretch receptive membrane of DB and after another time delay the membrane potential of DB begins to decrease. This decreases the drive to the muscle cells, which begin to relax. After another time delay the relaxation causes an increased signal from the stretch receptor, which increases the membrane potential of DB, and the cycle repeats.

The same interaction takes place between VB and the ventral muscle cells of the section. But these two oscillators are coupled by DD, VD and the forces exerted by one set of muscles on the other. This ensures that the two oscillators are 180 degrees out of phase with one another. Thus, when one set of muscle cells is maximally contracted, the other is maximally relaxed, which results in maximum flexion of the section.

There is also a time delay associated with the coupling between sections. Thus the instants of maximum flexion are displaced from one section to the next, resulting in the propagation of a wave of flexion down the length of the body of the worm.

Mechanical Model of the Organism

The curvature of the body is just the rate of change of the orientation with distance. Appendix C describes how a travelling wave arises from coupled oscillations along the

length of the organism. The orientation of the body as a function of distance from the tip of the head can be obtained by integrating equation (C.15). The exact solution of this problem involves an elliptic integal and is not of great importance to the present argument. However, an approximation of this function can be obtained by assuming that the curvature is constant throughout each of the ten sections. Then the posture of the animal can be approximated by summing the changes in orientation over all of the sections, as shown in figure 5.6. The resting length of each of the sections is L_R and the width, w, is the diameter of the body. These dimensions are constant by virtue of the structure of the cuticle, provided that the volume of the animal remains constant. Typically, L_R = 80 microns and w = 60 microns.

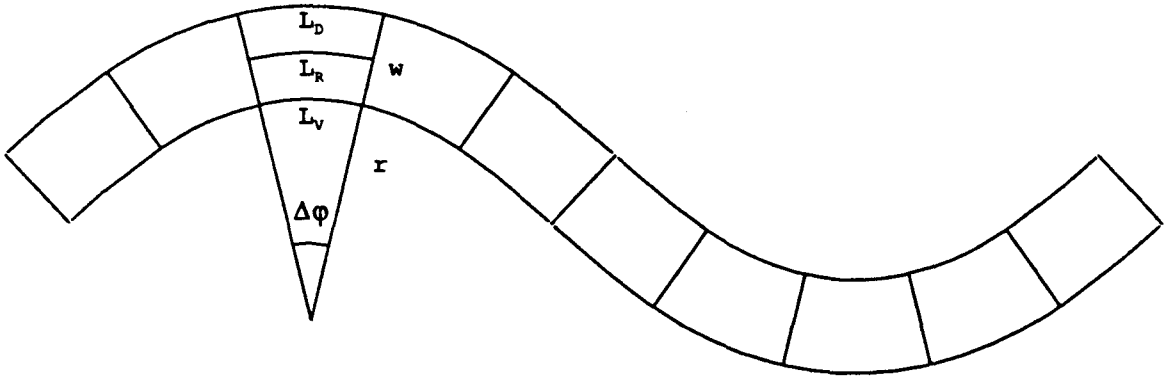

Figure 5.6 - Approximate posture of the body

If φ is the angle between the centre line of the body and the x axis and r is the radius of curvature of the ventral edge of the body, then the length of the ventral muscles is

given by

$$L_V = r\Delta\varphi \tag{18}$$

and the length of the dorsal muscles is

$$L_D = (r + w)\Delta\varphi \tag{19}$$

while the constant length of the centre line is

$$L_R = (r + \tfrac{1}{2}w)\Delta\varphi \tag{20}$$

When the section is straight, each of the muscles necessarily has the same length, namely L_R, so that L_R is the resting length of the muscles. The change in length resulting in the curvature $\Delta\varphi$ is just

$$L_D - L_R = \tfrac{1}{2}w\Delta\varphi \tag{21}$$

and

$$L_V - L_R = -\tfrac{1}{2}w\Delta\varphi \tag{22}$$

so that the shortening of one set of muscles is accompanied by an exactly equal lengthening of the opposing set. Denoting this change in length by ΔL,

$$\Delta\varphi = 2\frac{\Delta L}{w} \tag{23}$$

The change in length of the muscles results from the differential excitation received by the two sets of muscles from the motor nervous system.

A muscle may be modelled approximately by a spring with a variable force coefficient. The spring has a maximum and a minimum length, and the force constant is determined by the excitation to the muscle. In addition, any motion caused by changes in this excitement is damped by friction. This leads to the situation depicted in figure 5.7.

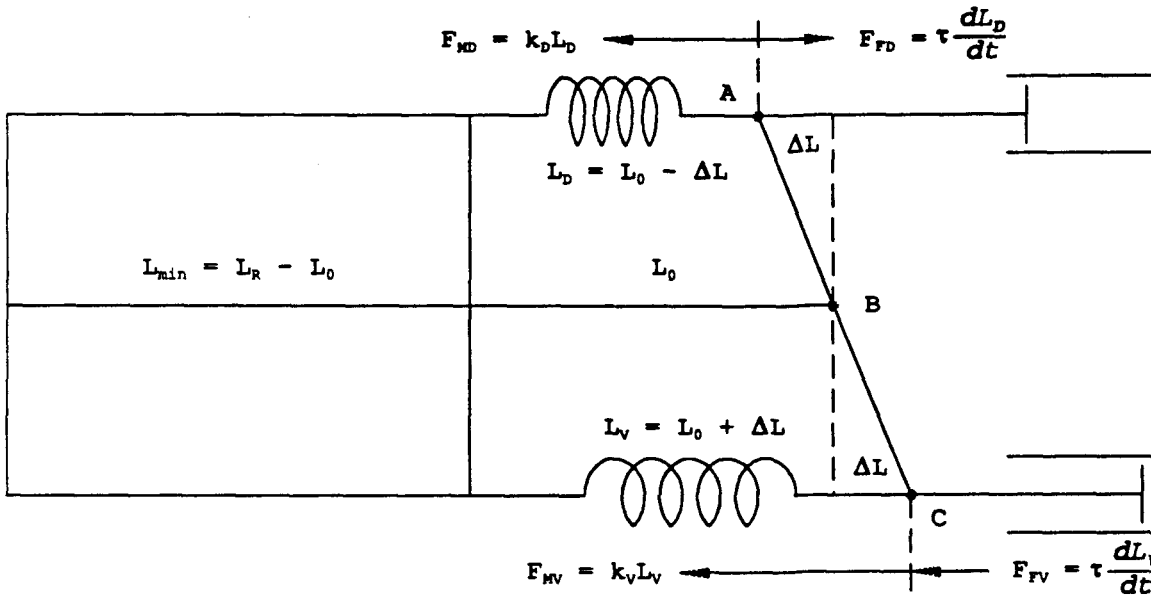

Figure 5.7 - Section geometry and force balance

In addition to the springs modelling the muscles, the frictional damping is represented by moving vanes whose motion is impeded by the viscosity of a fluid in which they move, and the constraint that the contraction of one set of muscles is equal to the stretch of the opposing set is introduced by the lever hinged at points A, B and C.

The net force exerted on the lever by the dorsal muscles is given by

$$F_D = F_{MD} + F_{FD} = k_D L_D + \tau \frac{dL_D}{dt} \tag{24}$$

and, substituting L = L - ΔL as depicted in the figure, which remains true even though the section may be curved, one obtains

$$F_D = k_D(L_0 - \Delta L) - \tau \frac{d}{dt}(\Delta L) \tag{25}$$

Similarly, the ventral muscles exert a net force of

$$F_V = k_V(L_0 + \Delta L) + \tau \frac{d}{dt}(\Delta L) \tag{26}$$

Now, the lever has a 1:1 ratio, hence the forces are reflected across the fulcrum unchanged in magnitude. Assuming that the motion of the system is dominated by the friction, so that the reaction to acceleration of the mass of the section is negligible due to the small mass and slow movements of the organism, the forces given by equations (25) and (26) must balance. Hence

$$F_D - F_V = 0 \tag{27}$$

That is,

$$(k_D - k_V) L_0 - (k_D + k_V) \Delta L - 2\tau \frac{d}{dt}(\Delta L) = 0 \tag{28}$$

and, defining x by

$$x_L = \frac{\Delta L}{L_0} \tag{29}$$

we have

$$2\tau \frac{dx_L}{dt} + (k_D + k_V) x_L = (k_D - k_V) \tag{30}$$

This equation provides the key ingredients necessary to change the idealized section of figure 5.5 into a model which can be analyzed mathematically in order to determine how the motor postprocessing system transforms a gating signal from the nerve ring into the precisely synchronized set of oscillating signals required to produce a wave of flexion travelling along the length of the body.

Modelling the Nerve Cords

The Boltzmann machine model presented becomes a useful substrate for modelling the nerve cords when it is modified to permit nontrivial time dependent behaviour. As it is typically described, the Boltzmann machine is gradually brought to thermodynamic equilibrium by a process of annealing. That is, the temperature is initially set to an effectively infinite value, and then reduced according to some fixed cooling schedule. As the temperature approaches zero, the machine freezes into some state representing the sought after mapping between input and output. The time dependent behaviour is thus fixed by the cooling schedule, which is

imposed from outside. This is clearly an insufficiently rich dynamics to account for the behaviour of the nerve cords.

However, letting the input signals themselves carry superimposed noise and giving each of the nodes a time constant over which the input signal is integrated permits a more general dependence on time. In this case, the temperature may be different for each node. The single global temperature is replaced by the equivalent noise temperature of the input signal at each node, as defined in equation (B.4). This results in a dynamics characterized by changes in the effective system structure via the mechanism of noise determined variable coupling between the nodes, as outlined in equations (B.8) and (B.12) and the related discussion. As well, this introduces the idea of a node as a lossy integrator driving a decision unit, which results in the nontrivial time behaviour described by the solutions of equation (B.28).

The first step in constructing a detailed mathematical model of the nerve cords is the identification of the state variables of the system, and the coupling between them. Not all of these state variables are associated with neural signals and not all of the coupling is synaptic. According to the paradigm which underlies this thesis, any force transmits information and any degree of freedom which is capable of exerting a force stores information. In particular, this applies to the differential length of the dorsal and ventral muscle bundles and to the forces due to the mechanical coupling between them. Any attempt to understand the

information processing which takes place in the nerve cords must therefore take this degree of freedom and these interactions into account.

As discussed above, each section can be decomposed into two subsystems, only one of which is active at any given time. These two subsystems are shown in figure 5.8, while three complete sections, including the coupling between them and the coupling common to all sections from the nerve ring, are shown in figure 5.9.

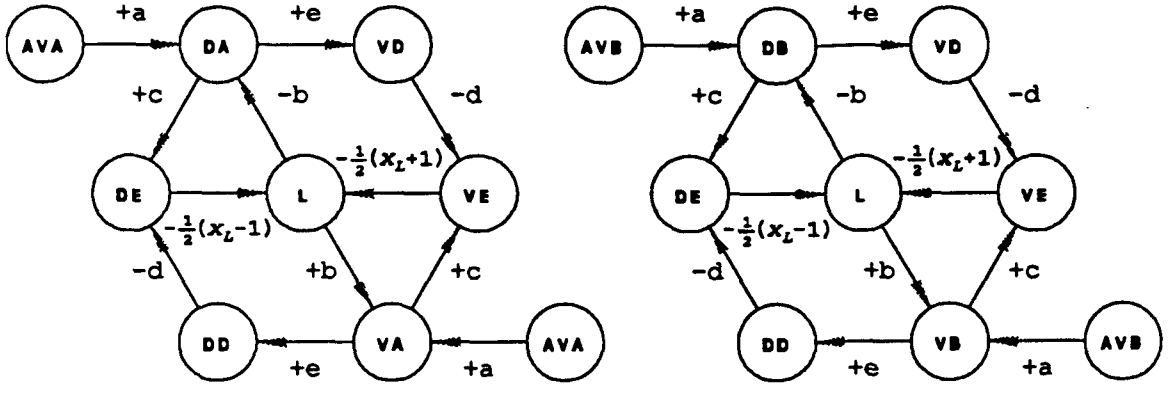

Figure 5.8 - The subsystems of a single section

In this figure the circles marked DE and VE indicate the excitation, x and x_v, to the dorsal and ventral muscles. The circles marked L indicate the state variable, x, associated with the difference in length between the two sets of muscles. Since this variable represents a stretch for the ventral muscles and a contraction for the dorsal muscles, the sign of the coupling coefficient is reversed for the dorsal stretch receptors. The coupling coefficients between the

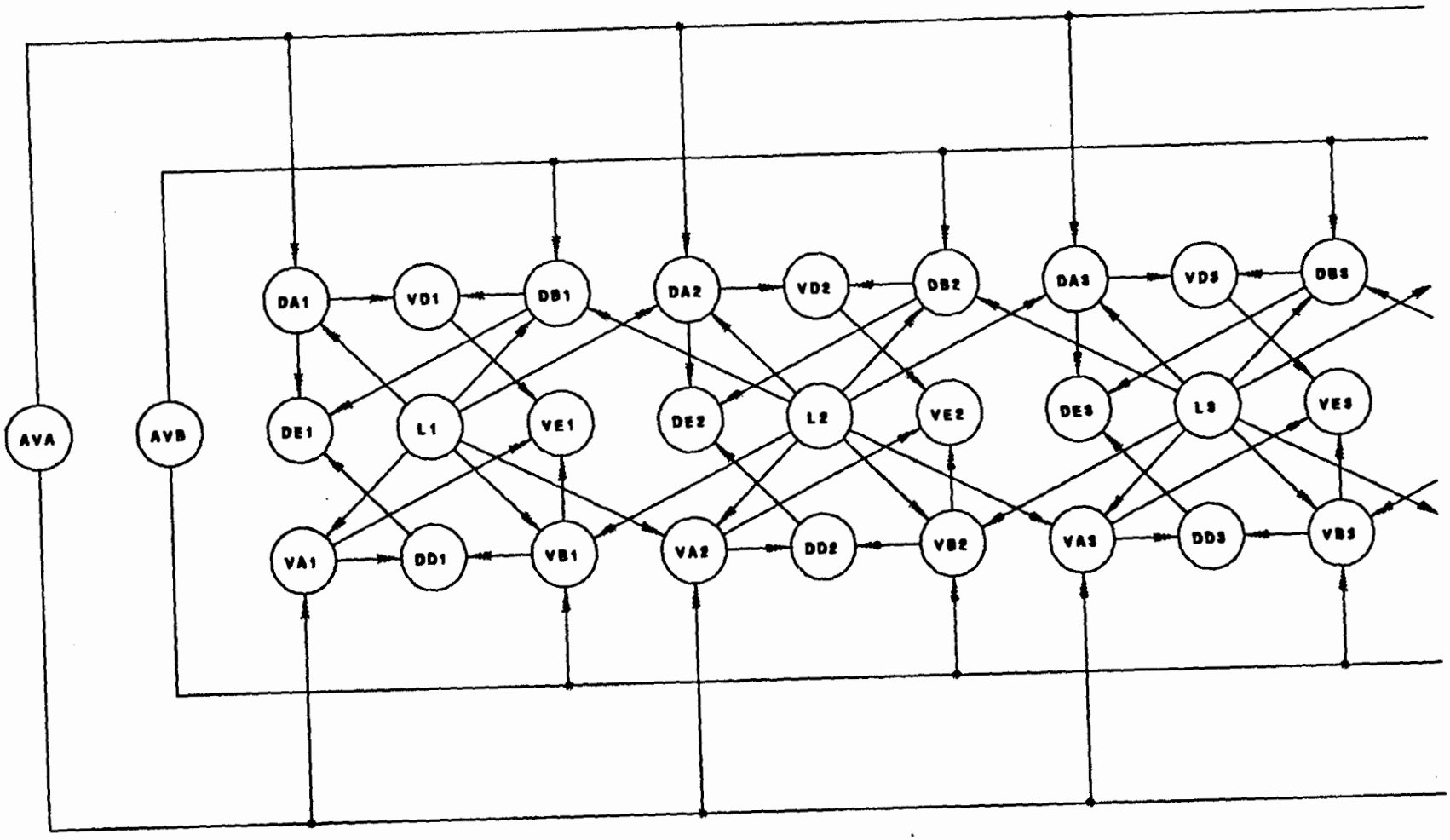

Figure 5.9 - The first three sections, showing inter-section coupling

excitations and the length state variable can be determined by rewriting equation (30) in the form

$$\tau \frac{dx_L}{dt} = -\frac{1}{2}(x_L - 1)k_D - \frac{1}{2}(x_L + 1)k_V \tag{31}$$

The two circles marked AVA represent the same signal; likewise for the two circles marked AVB. These are the gating signals from the nerve ring and are applied to every section simultaneously. Travelling waves are produced only when one of these signals is on an the other one is off. When both are on, two waves travelling in opposite directions are generated and superpose to produce a resultant standing wave. This results in the organism assuming some fixed posture consisting of a stationary wave of flexion along the length of its body. As will be shown below, these subsystems exhibit stable oscillations only when biased appropriately by their gating signals.

One additional assumption must be made. Namely, that the response of a muscle cell to excitation can be expressed by the same sigmoid transfer function that governs the behaviour of neurons. That is,

$$k_D = k\,s(x_{DE}), \quad k_V = k\,s(x_{VE}) \tag{32}$$

where s is the sigmoid, given by equation (3) and x and x_v are the effective membrane potentials of the muscle cells. Recall that the effective membrane potentials are short term time averages of stochastic signals, characterized by a slowly

varying mean value about which the variations oscillate, as defined by equation (12) and as discussed in greater detail in appendix B.

It must also be noted that under steady state conditions, the coefficient of x in equation (30) is constant. This is because the excitation to dorsoventrally opposed muscles varies in antiphase, so that

$$s(x_{DE}) + s(x_{VE}) = f, \quad s(x_{DE}) - s(x_{VE}) = g(t) \tag{33}$$

for some positive constant, f, and periodic function, g. Hence

$$\frac{2\tau}{f}\frac{dx_L}{dt} = -x_L + \frac{1}{f}s(x_{DE}) - \frac{1}{f}s(x_{VE}) \tag{34}$$

where the constant, k, has been absorbed into f.

From figure 5.8 one can then immediately write the differential equations of motion of the subsystem which mediates forward motion. Letting τ be the membrane time constant of the neurons, τ be the membrane time constant of the muscle cells and τ be the somatic time constant due to friction in the body section, one obtains

$$\begin{aligned}
\tau_n \dot{x}_{VD} &= -x_{VD} + e\,s(x_{DB}) \\
\tau_n \dot{x}_{DB} &= -x_{DB} - b\,x_L + a\,s(x_{AVD}) \\
\tau_m \dot{x}_{DE} &= -x_{DE} + c\,s(x_{DB}) - d\,s(x_{DD}) \\
\frac{2}{f}\tau_s \dot{x}_L &= -x_L + \frac{1}{f}s(x_{DE}) - \frac{1}{f}s(x_{VE}) \\
\tau_m \dot{x}_{VE} &= -x_{VE} + c\,s(x_{VB} - d\,s(x_{VD}) \\
\tau_n \dot{x}_{VB} &= -x_{VB} + b\,x_L + a\,s(x_{AVD}) \\
\tau_n \dot{x}_{DD} &= -x_{DD} + e\,s(x_{VB})
\end{aligned} \tag{35}$$

In order to simplify the analysis, the membrane time constants of the neurons and muscle cells may be taken as approximately equal. The somatic time constant will undoubtedly by somewhat longer. It is not unreasonable, then, to assume

$$\tau_n = \tau_m = \frac{2}{f}\tau_s = \tau \tag{36}$$

This assumption will facilitate the calculation of the solution without affecting the nature of the solution. We may now write equations (35) explicitly as a vector differential equation:

$$\tau\frac{d}{dt}\begin{bmatrix} x_{VD} \\ x_{DB} \\ x_{DE} \\ x_L \\ x_{VE} \\ x_{VB} \\ x_{DD} \end{bmatrix} + \begin{bmatrix} x_{VD} \\ x_{DB} \\ x_{DE} \\ x_L \\ x_{VE} \\ x_{VB} \\ x_{DD} \end{bmatrix} = \begin{bmatrix} \cdot & -e & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & b & \cdot & \cdot & \cdot \\ \cdot & -c & \cdot & \cdot & \cdot & \cdot & d \\ \cdot & \cdot & -\frac{1}{f} & \cdot & \frac{1}{f} & \cdot & \cdot \\ d & \cdot & \cdot & \cdot & \cdot & -c & \cdot \\ \cdot & \cdot & \cdot & -b & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & -e & \cdot \end{bmatrix} \begin{bmatrix} s(x_{VD}) \\ s(x_{DB}) \\ s(x_{DE}) \\ x_L \\ s(x_{VE}) \\ s(x_{VB}) \\ s(x_{DD}) \end{bmatrix} + \begin{bmatrix} \cdot \\ a \\ \cdot \\ \cdot \\ \cdot \\ a \\ \cdot \end{bmatrix} \tag{37}$$

where the matrix on the right hand side is just M, the coupling matrix, augmented to take into account coupling introduced by sensory and mechanical interactions, with the condition expressed by equation (31) imposed. As well, $s(x_v)$ has been absorbed into the coefficient, a, in order to simplify the representation of excitation from the nerve ring. This is an example of the noise determined variable coupling described in appendix B. Forming the interaction matrix, N, and absorbing the slopes of the sigmoids into the coefficients gives

$$\tau\frac{d}{dt}\begin{bmatrix} x_{VD} \\ x_{DB} \\ x_{DE} \\ x_{L} \\ x_{VE} \\ x_{VB} \\ x_{DD} \end{bmatrix} + \begin{bmatrix} 1 & -e & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & b & \cdot & \cdot & \cdot \\ \cdot & -c & 1 & \cdot & \cdot & \cdot & d \\ \cdot & \cdot & -\frac{1}{f} & 1 & \frac{1}{f} & \cdot & \cdot \\ d & \cdot & \cdot & \cdot & 1 & -c & \cdot \\ \cdot & \cdot & \cdot & -b & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & -e & 1 \end{bmatrix} \begin{bmatrix} x_{VD} \\ x_{DB} \\ x_{DE} \\ x_{L} \\ x_{VE} \\ x_{VB} \\ x_{DD} \end{bmatrix} = \begin{bmatrix} \cdot \\ a \\ \cdot \\ \cdot \\ \cdot \\ a \\ \cdot \end{bmatrix} \tag{38}$$

where the matrix on the left hand side is just A = (I - N). This system can be expressed in terms of the sums and differences of the homologous dorsoventrally opposed signals as outlined in appendix D. One obtains

$$\tau\frac{d}{dt}\begin{bmatrix} x_{DM} \\ x_{BM} \\ x_{EM} \\ x_{L} \\ x_{EV} \\ x_{BV} \\ x_{DV} \end{bmatrix} + \begin{bmatrix} 1 & -e & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ d & -c & 1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 & \frac{2}{f} & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 & -c & -d \\ \cdot & \cdot & \cdot & -b & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & -e & 1 \end{bmatrix} \begin{bmatrix} x_{DM} \\ x_{BM} \\ x_{EM} \\ x_{L} \\ x_{EV} \\ x_{BV} \\ x_{DV} \end{bmatrix} = \begin{bmatrix} \cdot \\ a \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} \tag{39}$$

where the $x_{\cdot}$, being the sums, represent the mean values and the $x_{\cdot V}$, being the differences, represent the variations about the means. Now, equation (39) is highly revealing of the structure of the subsystem, because it clearly shows it to be composed of two smaller subsystems which are entirely uncoupled. The time development of the means is completely decoupled from the time development of the variations. We may therefore proceed to solve equation (39) in two stages. Each of the two reduced subsystems will be treated separately and the solutions will then be combined. The details of these solutions are provided in appendix D, wherein the structure of

the eigenfunctions and the generalized eigenvectors of the two subsystems are explicitly derived.

Time Evolution of the Means

The solution of the differential equation of motion of the means, expresssed explicitly in terms of the parameters of the system, is given by

$$\begin{bmatrix} x_{DM} \\ x_{BM} \\ x_{EM} \end{bmatrix} = a\left[\begin{bmatrix} e \\ 1 \\ c-de \end{bmatrix} + \begin{bmatrix} e\left(\frac{t}{\tau}-1\right) \\ -1 \\ (c-de)\left(\frac{t}{\tau}-1\right)+\frac{de}{2}\frac{t^2}{\tau^2} \end{bmatrix}\exp\left(-\frac{t}{\tau}\right)\right] \tag{40}$$

Thus, η is composed of a constant part and a transient part. It is clear by inspection of this equation that η starts from zero and asymptotically approaches

$$\eta_{\infty} = a\begin{bmatrix} e \\ 1 \\ c-de \end{bmatrix} \tag{41}$$

as t increases without bound. Thus the assumption expressed in equation (33) is validated, with

$$f = 2s(a(c-de)) \tag{42}$$

This argument is not circular because the operations leading to equation (39) are independent of f. It is also clear from equation (40) that $\eta = 0$ for any value of t in the absence of excitation from the nerve ring.

The time evolution of the means is shown in figure 5.10.

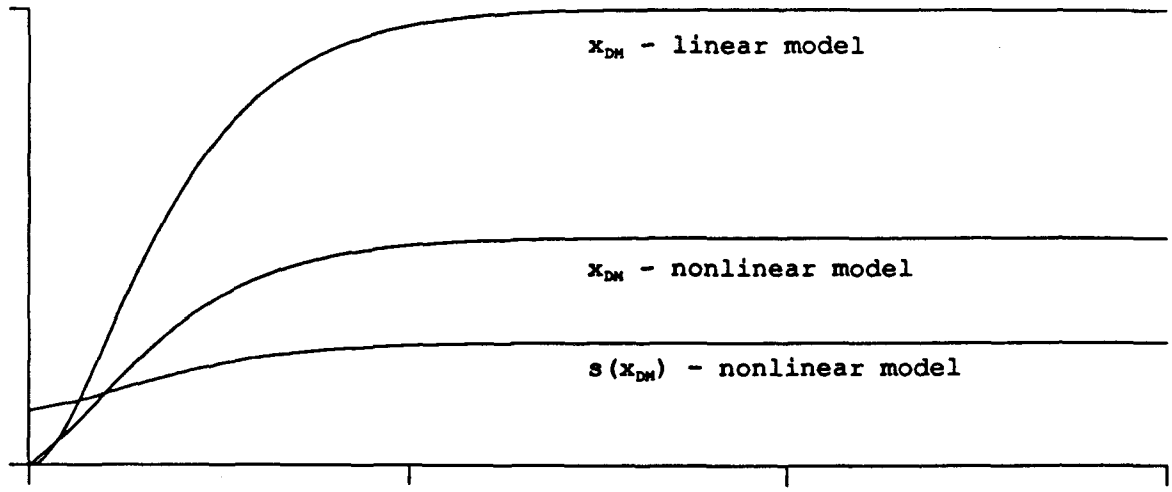

Figure 5.10 a - Mean time response of the DD and VD neurons

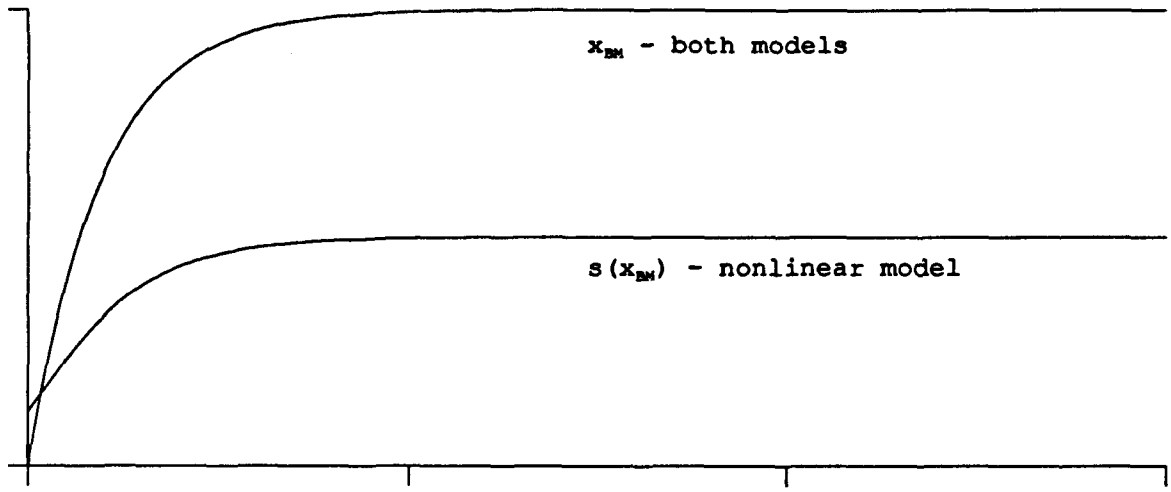

Figure 5.10 b - Mean time response of the DB and VB neurons

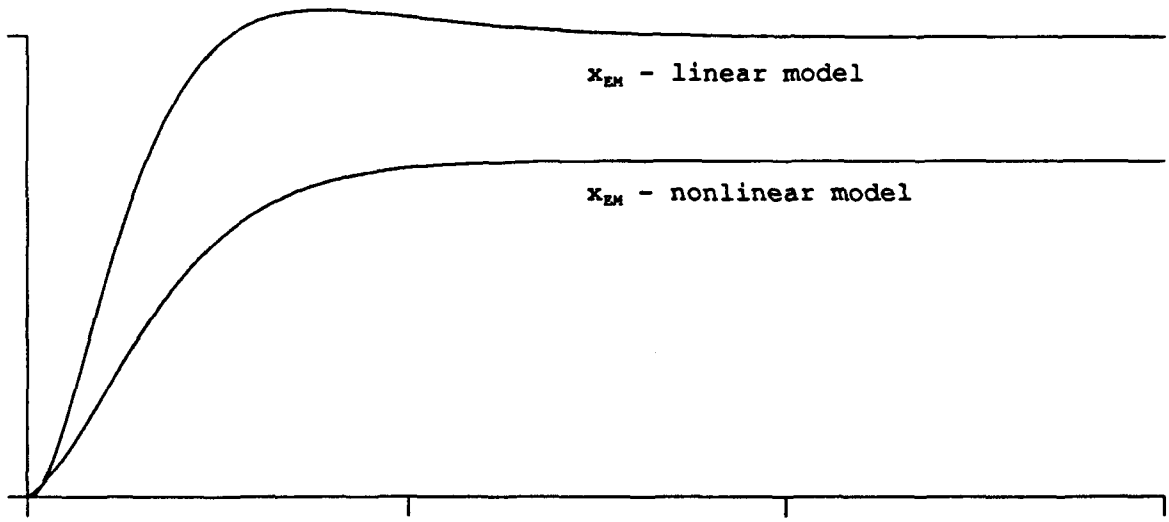

Figure 5.10 c - Mean time response of the muscular excitations

Simulations were carried out with a = 1, b = 5, c = 2, d = 2, e = 1 and f = 1. Both linearized and nonlinear models were run. In the figure, the vertical scale is one unit per division, with one unit representing the maximum output from a neuron. The horizontal scale is five time constants per division. As can be seen in the figure, the form of the solutions is identical, or virtually so, between the linear and nonlinear models. The only significant difference between the solutions to the two models is the value of the asymptotes.

Time Evolution of the Variations

Turning now to the variation, η_V, about the mean, η, equation (40) reduces to

$$\tau\frac{d}{dt}\begin{bmatrix} x_L \\ x_{EV} \\ x_{BV} \\ x_{DV} \end{bmatrix} + \begin{bmatrix} 1 & \frac{2}{f} & \cdot & \cdot \\ \cdot & 1 & -c & -d \\ -b & \cdot & 1 & \cdot \\ \cdot & \cdot & e & 1 \end{bmatrix}\begin{bmatrix} x_L \\ x_{EV} \\ x_{BV} \\ x_{DV} \end{bmatrix} = \begin{bmatrix} \cdot \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} \tag{43}$$

or just

$$t\frac{d}{dt}\eta_V + A_V\eta_V = 0 \tag{44}$$

where, it will be recalled, the slopes of the sigmoids are absorbed into the coefficients. These slopes, of course, are to be evaluated at the mean, η. Hence, by equation (41), excitation from the nerve ring is required in order for equation (44) to have a nontrivial solution. In the absence

of such excitation, η decays to zero, the system is biased into the region where slopes of the sigmoids are nearly zero, and the system becomes effectively decoupled. When the nerve ring drives the system into the central regions of the sigmoids, small random fluctuations are able to initiate the oscillatory dynamics which lead to motion of the organism.

The detailed solution of equation (43) is given in appendix D. Assuming that the initial conditions are such that only the initial posture of the organism contributes to the solution, the time behaviour of η_v satisfies

$$\eta_v = 2\left[\begin{bmatrix} r_1^4\cos(\omega_1 t+\theta_3-4\theta_1 \\ \frac{f}{2}r_1^5\cos(\omega_1 t+\theta_3-5\theta_1 \\ -br_1^3\cos(\omega_1 t+\theta_3-3\theta_1 \\ ber_1^2\cos(\omega_1 t+\theta_3-2\theta_1 \end{bmatrix} r_3 e^{-(a+1)\frac{t}{\tau}} + \begin{bmatrix} r_2^4\cos(\omega_2 t+\theta_4+4\theta_2) \\ r_2^5\cos(\omega_2 t+\theta_4+5\theta_2) \\ r_2^3\cos(\omega_2 t+\theta_4+3\theta_2) \\ r_2^2\cos(\omega_2 t+\theta_4+2\theta_2) \end{bmatrix} r_4 e^{(a-1)\frac{t}{\tau}}\right] x_{L0} \qquad (45)$$

Now, the first term on the right hand side of this equation is exponentially damped, so it will decay to zero over time. The condition for the second term to grow exponentially with time is examined more closely in appendix D.

Of course, the oscillation will not grow in magnitude without bound as indicated by equation (45). Eventually, the linearizing assumptions will break down as the amplitude drives the peaks of the signal into the regions where the slopes of the sigmoids decrease enough to have a significant effect on the interaction matrix. The result will be to decrease all of the coupling coefficients except f. The oscillations will stabilize when the effective value of the

left hand side of inequality (D.53), integrated over the whole signal swing, reaches unity. In particular, the effective frequency of oscillation will be lowered. The peaks of the waveforms will also be flattened as the signal excursion takes it into regions in which the slopes of the sigmoids are significantly reduced.

These effects can be seen in figures 5.11, 5.12 and 5.13. In figure 5.11, the oscillatory components of the signals are shown for the first 15 time constants after the arrival of a gating signal from the nerve ring. Both the membrane potentials and the outputs are shown for the neural signals. For comparison, the response of the linearized system is shown in figure 5.12. The horizontal scale is the same in both figures, namely, five time constants per division. The vertical scale in figure 5.11 is normalized to the maximum output of a neuron. The vertical scale in figure 5.12 is 30 times this amount.

In figure 5.13, the time evolution of the system is traced for 15 time constants following the arrival of the gating signal. This interval is of sufficient length for the system to reach a steady oscillatory state. Note how, even though the peaks of the outputs of the DB and VB neurons are flattened, the resulting differential muscle length remains nearly sinusoidal. This was an unexpected feature of the system dynamics and bears further investigation. One possible explanation is that the coupling coefficient, f, between the cord motor neurons and the muscles decreases as the signals

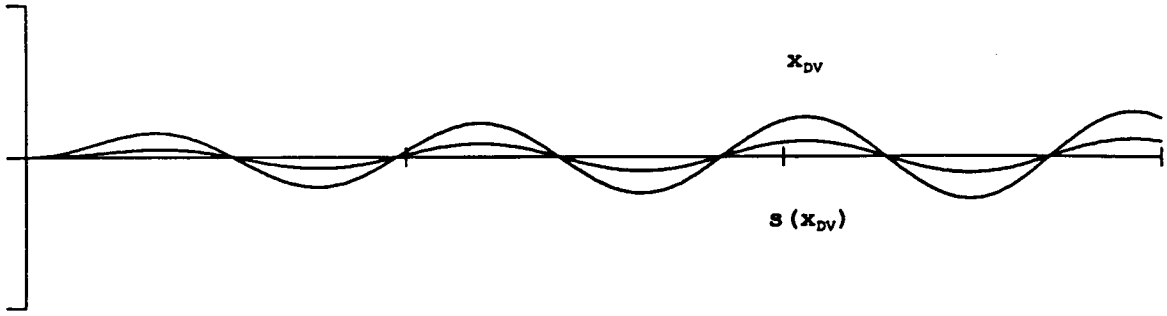

Figure 5.11 a - Differential time response of the DD and VD neurons

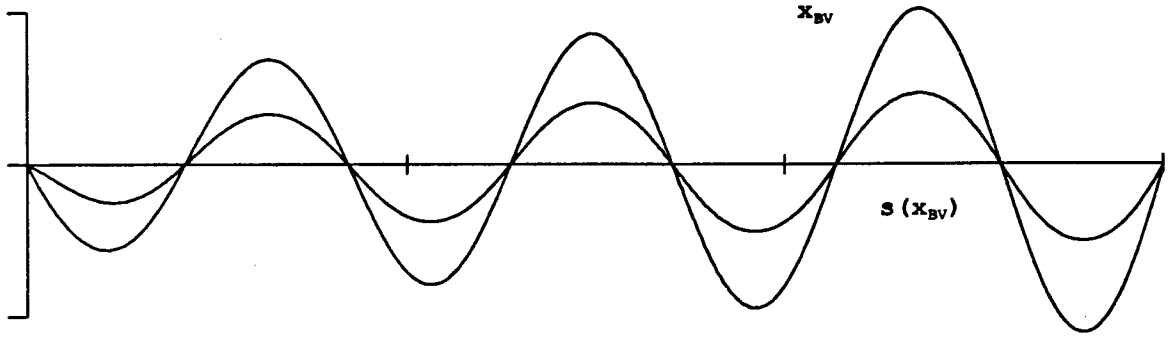

Figure 5.11 b - Differential time response of the DB and VB neurons

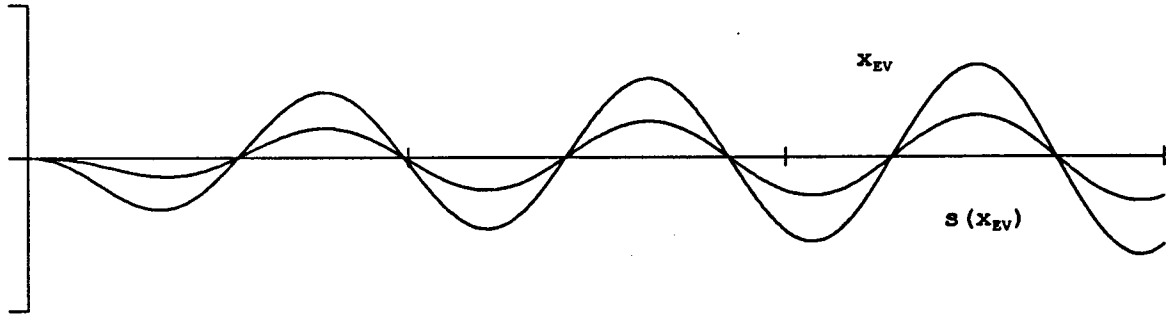

Figure 5.11 c - Differential time response of the muscular excitations

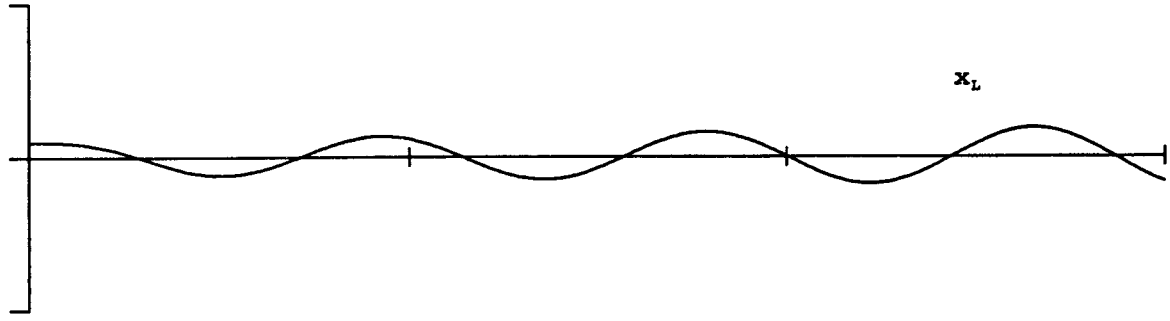

Figure 5.11 d - Time evolution of the differential muscle lengths

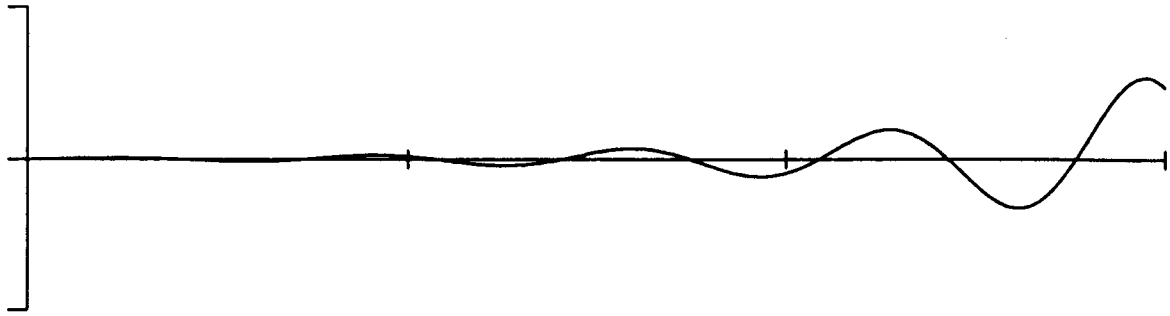

Figure 5.12 a - Time evolution of x_{DV} in the linear model

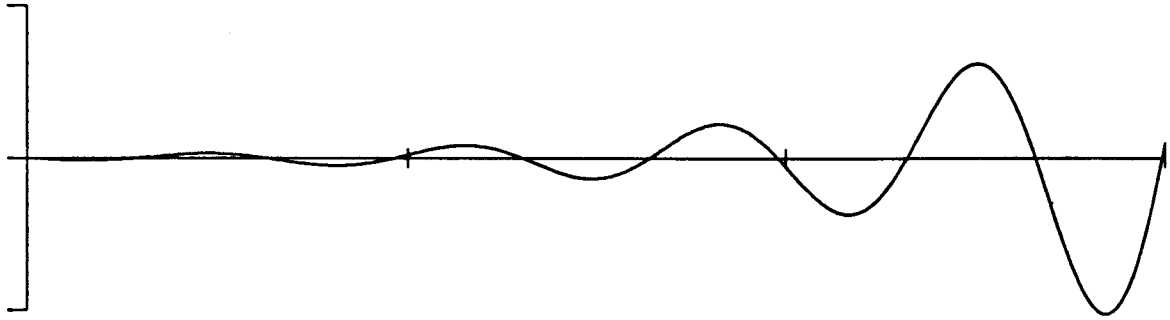

Figure 5.12 b - Time evolution of x_{BV} in the linear model

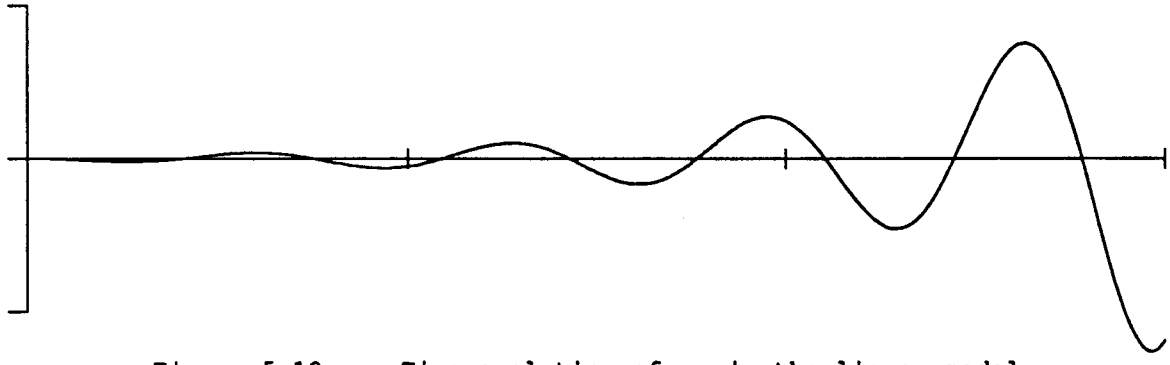

Figure 5.12 c - Time evolution of x_{EV} in the linear model

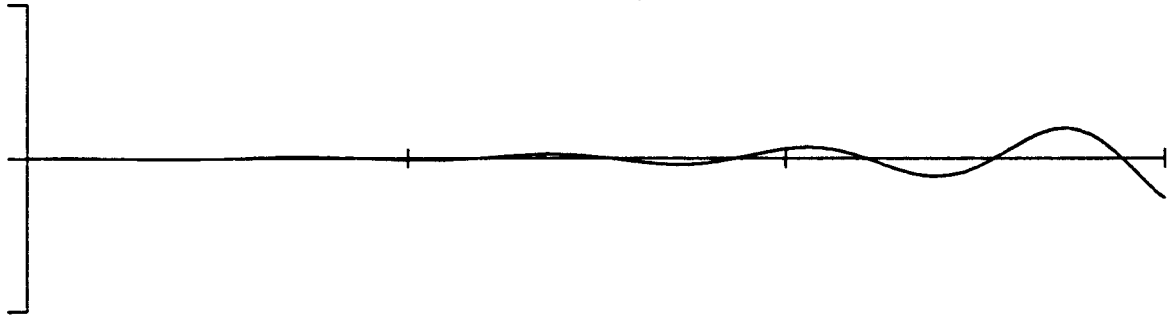

Figure 5.12 d - Time evolution of x_L in the linear model

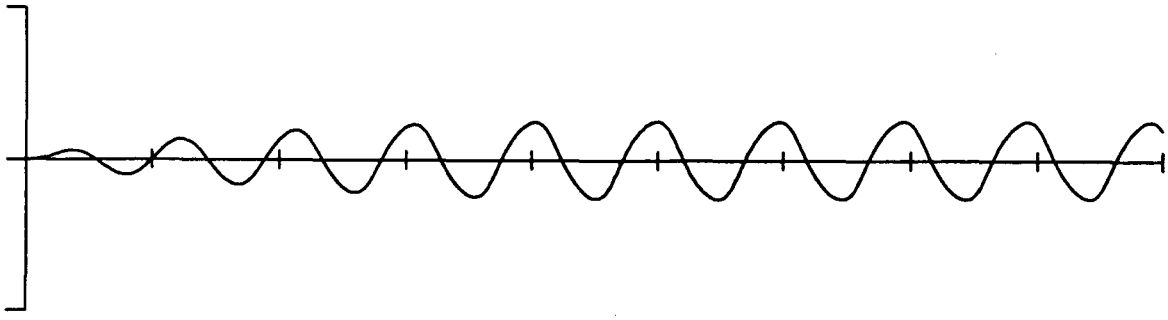

Figure 5.13 a - Differential time response of the DD and VD neurons

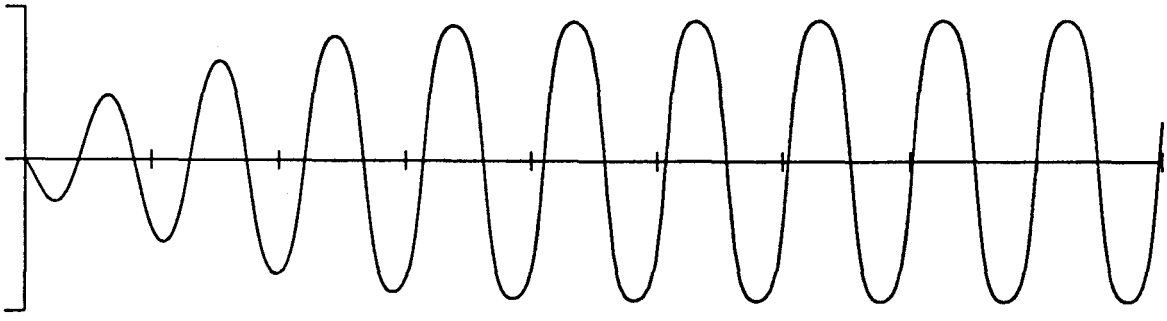

Figure 5.13 b - Differential time response of the DB and VB neurons

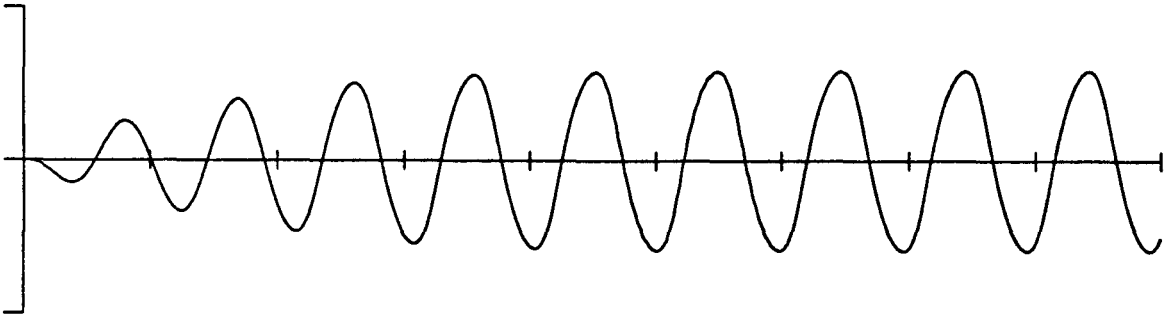

Figure 5.13 c - Differential time response of the muscular excitations

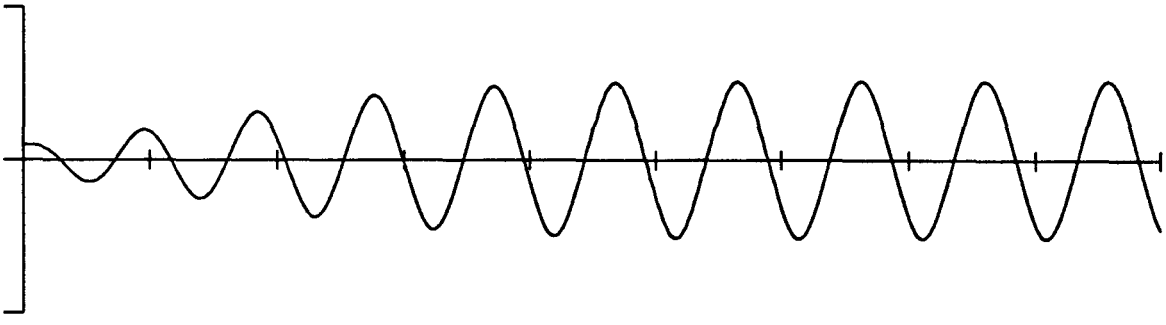

Figure 5.13 d - Time evolution of the differential muscle lengths

approach the limits of their excursions. It is also worthy of note that the waveforms of the other neural signal takes on a slight sawtooth character as the system enters the steady state.

Overall Time Response of One Section

With the solutions now in hand both for the time development of the mean values of the state variables and for the variations about the means, it becomes possible to construct the overall solution to equation (38) and to compare it to the results obtained by simulating the fully nonlinear system expressed by equation (35). Under steady state conditions, the first term on the right hand side of equation (45) will have decayed to negligible amplitude, as will the second term on the right hand side of equation (40). Thus

$$\begin{bmatrix} x_{DN} \\ x_{BM} \\ x_{EM} \\ x_{L} \\ x_{EV} \\ x_{BV} \\ x_{DV} \end{bmatrix} = \begin{bmatrix} ae \\ a \\ a(c-de) \\ 2x_{LO}r_2^4 r_4 \cos(\omega_2 t + \theta_4 + 4\theta_2) \\ -x_{LO} f r_2^5 r_4 \cos(\omega_2 t + \theta_4 + 5\theta_2) \\ 2x_{LO} b r_2^3 r_4 \cos(\omega_2 t + \theta_4 + 3\theta_2) \\ 2x_{LO} be r_2^2 r_4 \cos(\omega_2 t + \theta_4 + 2\theta_2) \end{bmatrix} \tag{46}$$

where all constants that depend on the coupling coefficients are evaluated using the effective values of these coefficients when integrated over the signal swing. Thus, equation (46) is, strictly speaking, only the first term of the Fourier expansion of the actual signal.

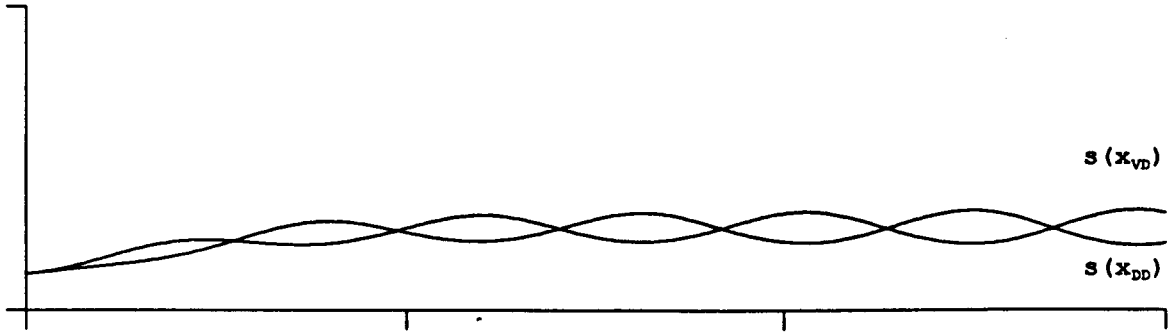

Figure 5.14 a - Overall time response of the DD and VD neurons

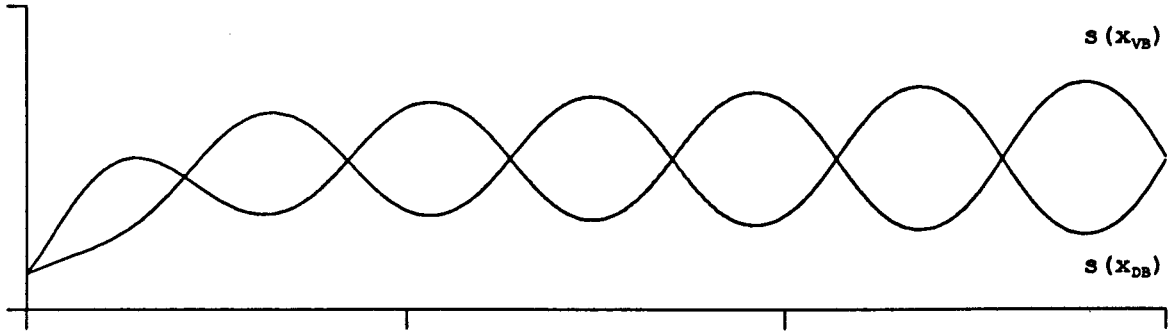

Figure 5.14 b - Overall time response of the DB and VB neurons

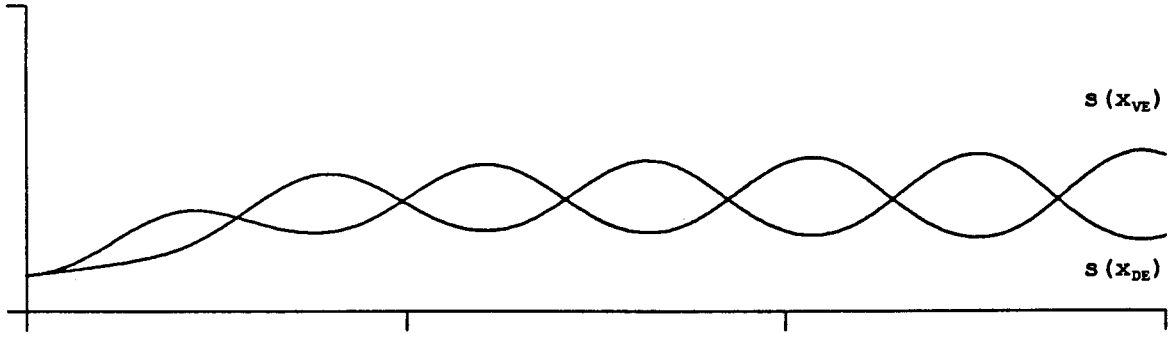

Figure 5.14 c - Overall time response of the muscular excitations

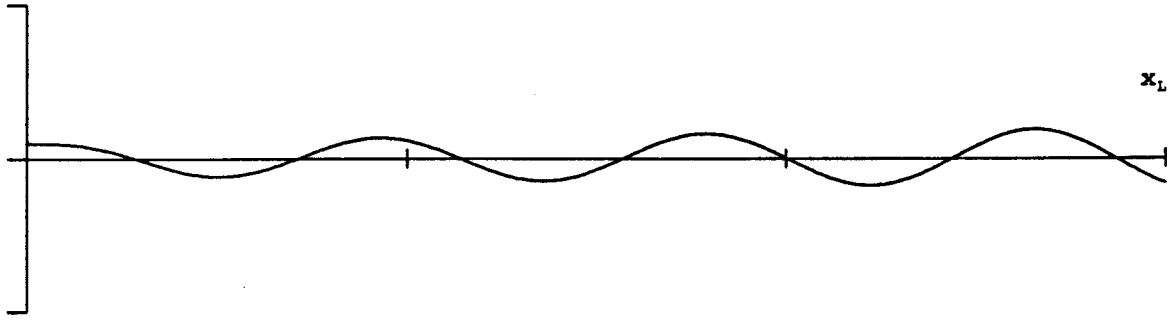

Figure 5.14 d - Time evolution of the differential muscle lengths

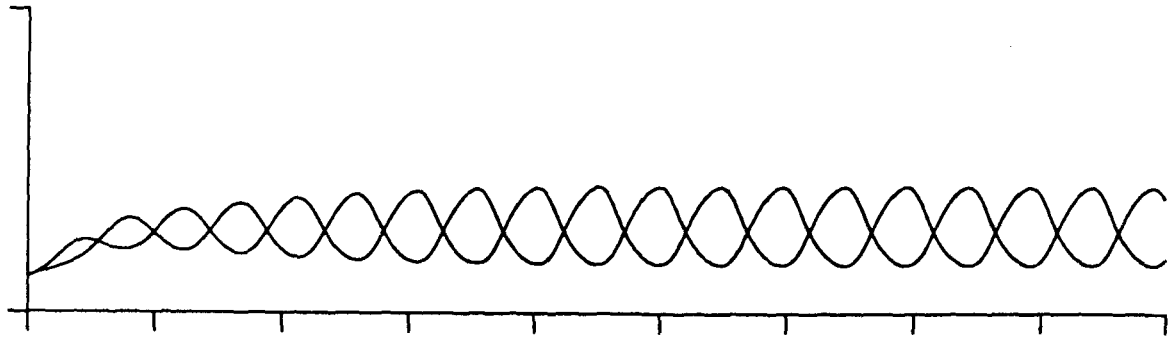

Figure 5.15 a - Overall time response of the DD and VD neurons

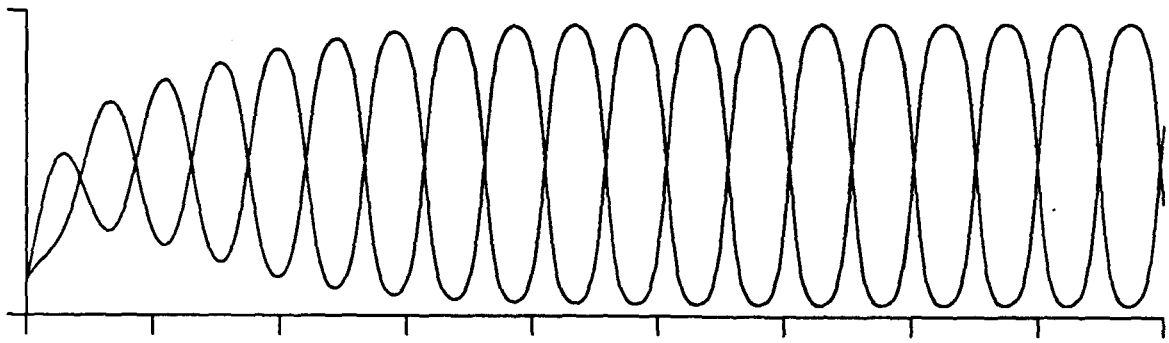

Figure 5.15 b - Overall time response of the DB and VB neurons

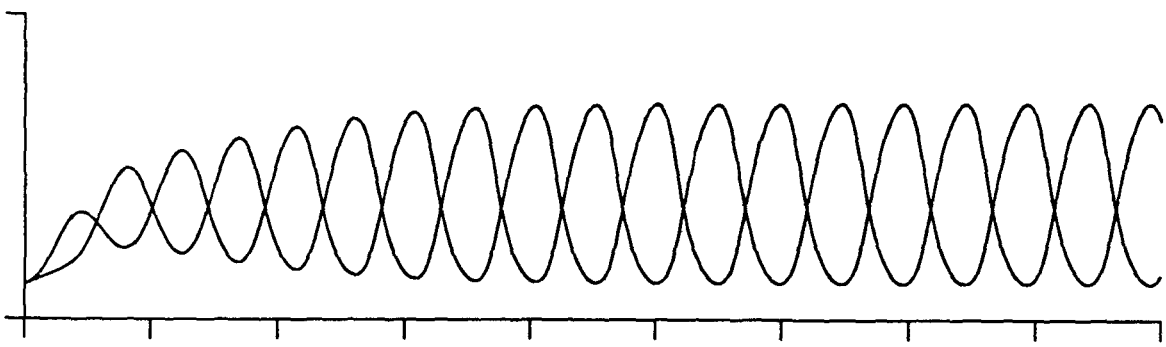

Figure 5.15 c - Overall time response of the muscular excitations

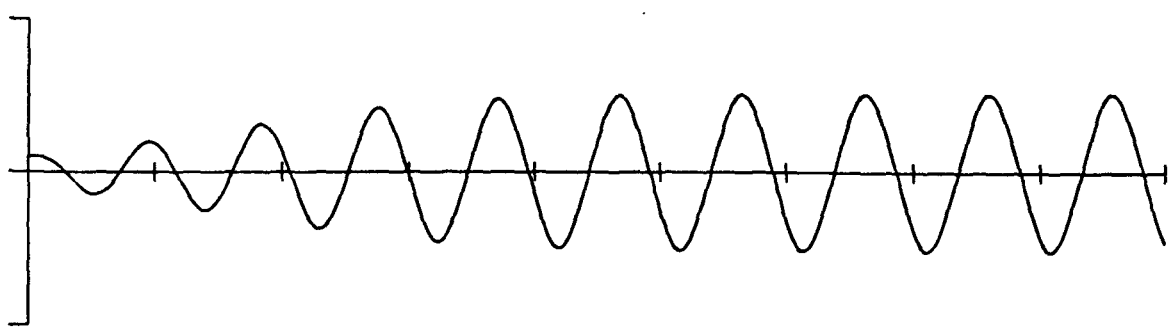

Figure 5.15 d - Time evolution of the differential muscle lengths

Finally, transforming back to the actual membrane potentials from the sums and differences of equation (46) gives

$$\begin{bmatrix} x_{VD} \\ x_{DB} \\ x_{DE} \\ x_{L} \\ x_{VE} \\ x_{VB} \\ x_{DD} \end{bmatrix} = \begin{bmatrix} ae + 2x_{L0}ber_2^2r_4\cos(\omega_2 t + \theta_4 + 2\theta_2) \\ a + 2x_{L0}br_2^3r_4\cos(\omega_2 t + \theta_4 + 3\theta_2) \\ a(c-de) - x_{L0}fr_2^5r_4\cos(\omega_2 + \theta_4 + 5\theta_2) \\ 2x_{L0}r_2^4r_4\cos(\omega_2 t + \theta_4 + 4\theta_2) \\ a(c-de) + x_{L0}fr_2^5r_4\cos(\omega_2 t + \theta_4 + 5\theta_2) \\ a - 2x_{L0}br_2^3r_4\cos(\omega_2 t + \theta_4 + 3\theta_2) \\ ae - 2x_{L0}ber_2^2r_4\cos(\omega_2 t + \theta_4 + 2\theta_2) \end{bmatrix} \tag{47}$$

Hence the dorsoventrally opposed signals oscillate in anti-phase about identical mean values when the system is gated by the nerve ring. Otherwise the activity decays to zero. These signals are shown in figures 5.14 and 5.15. For ease of comparison with figures 5.10, 5.11 and 5.13, figure 5.14 traces the evolution for five time constants while figure 5.15 covers 15 time constants. An interesting effect that appears in figure 5.15 which was not apparent in figure 5.13 is the distinct asymmetry in the signal excursions of DD, VD and the muscular excitation. Again, the likely explanation is the systematic variation of f with signal amplitude.

Overall Time Response of the Nerve Cords

It now remains to take into account the effect of the inter-section coupling due to the distal stretch receptive endings of the DB and VB neurons. The effect of this coupling is to enforce a phase constraint between the sections which

gives rise to a wave of flexion which travels along the body of the organism at a constant rate.

If the coupled signal, x_{BV}, has the same phase as the unperturbed signal, x_{BVO}, due only to the internal interactions within the section, it will tend to increase the amplitude of the resultant, x_{BVR}, beyond that implied by inequality (D.53). As a result, the exponent in the second term on the right hand side of equation (45) will become negative and the amplitude will tend to decay. A new equilibrium will be established when these two effects balance, in which case the resultant is again a sinusoid of constant amplitude. Denoting this amplitude by A and, for the sake of simplicity, choosing the time origin such that $\theta + 3\theta = 0$, this equilibrium is expressed by

$$\tau \frac{d}{dt} x_{BVR} = \tau \frac{d}{dt} x_{BVO} + x_{BVC} = \tau \frac{d}{dt} A \cos \omega_2 t \tag{48}$$

but, from equation (45) one has

$$\tau \frac{d}{dt} x_{BVO} = -A\omega_2 \tau \sin \omega_2 t + (\alpha - 1) \cos \omega_2 t \tag{49}$$

and the first term on the right hand side of this equation is just

$$-A\omega_2 \tau \sin \omega_2 t = \tau \frac{d}{dt} A \cos \omega_2 t \tag{50}$$

The requirement for x_{BV} to have the same phase as x_{BVO} implies

that

$$x_{BVC} = B\cos\omega_2 t \tag{51}$$

for some amplitude, B. Combining these four expressions yields

$$(\alpha - 1)A = B \tag{52}$$

as the new condition for equilibrium. By driving the section into regions in which inequality (D.53) is no longer satisfied, the coupling causes the section to adjust its amplitude and phase until equation (52) is satisfied.

Now, the distal stretch receptive endings are located in the next section and their signal must therefore travel twice as far, on the average, as the signal from the proximal stretch receptive segments providing feedback within the section. This corresponds to a phase shift of θ. Hence, moving the time origin again to allow the stretch sensing signal from the next section to be written

$$x_{BVR}^{(n+1)} = 2x_{LO}^{(n+1)} b r_2^3 r_4 \cos\omega_2 t \tag{53}$$

one obtains

$$x_{BVC}^{(n+1)} = 2x_{LO}^{(n+1)} b r_2^3 r_4 \cos(\omega_2 t + \theta_2) \tag{54}$$

and, recalling that

$$A = 2x_{LO}^{(n)} b r_2^3 r_4 \tag{55}$$

equation (52) becomes

$$(\alpha - 1) = \frac{x_{LO}^{(n+1)}}{x_{LO}^{(n)}} \tag{56}$$

and, adjusted for the new time origin, equation (48) implies that

$$x_{BVR}^{(n)} = 2x_{LO}^{(n)} b r_2^3 r_4 \cos(\omega_2 t + \theta_2) \tag{57}$$

Hence the sections oscillate at the same frequency, ω, and at a constant phase angle, θ, from each other. This phase shift occurs over the length, L_R, of one section. Hence, by equation (C.12), the velocity at which the wave of flexion travels down the body of the organism is

$$v = -\frac{\omega_2 L_R}{\theta_2} \tag{58}$$

under steady state conditions.

Figure 5.16 shows the differential muscle lengths of the ten sections of somatic musculature when the subsystem composed of DB, DD, VB and VD is gated in each section. All sections are gated simultaneously by the nerve ring. In this simulation, only the last section was given a nonzero initial curvature. As can be seen in the figure, the oscillations begin in that section and propagate forward along the length of the organism until all sections are oscillating. By the end of 15 time constants, most of the sections have reached a steady state condition.

The model presented in this section thus exhibits all of

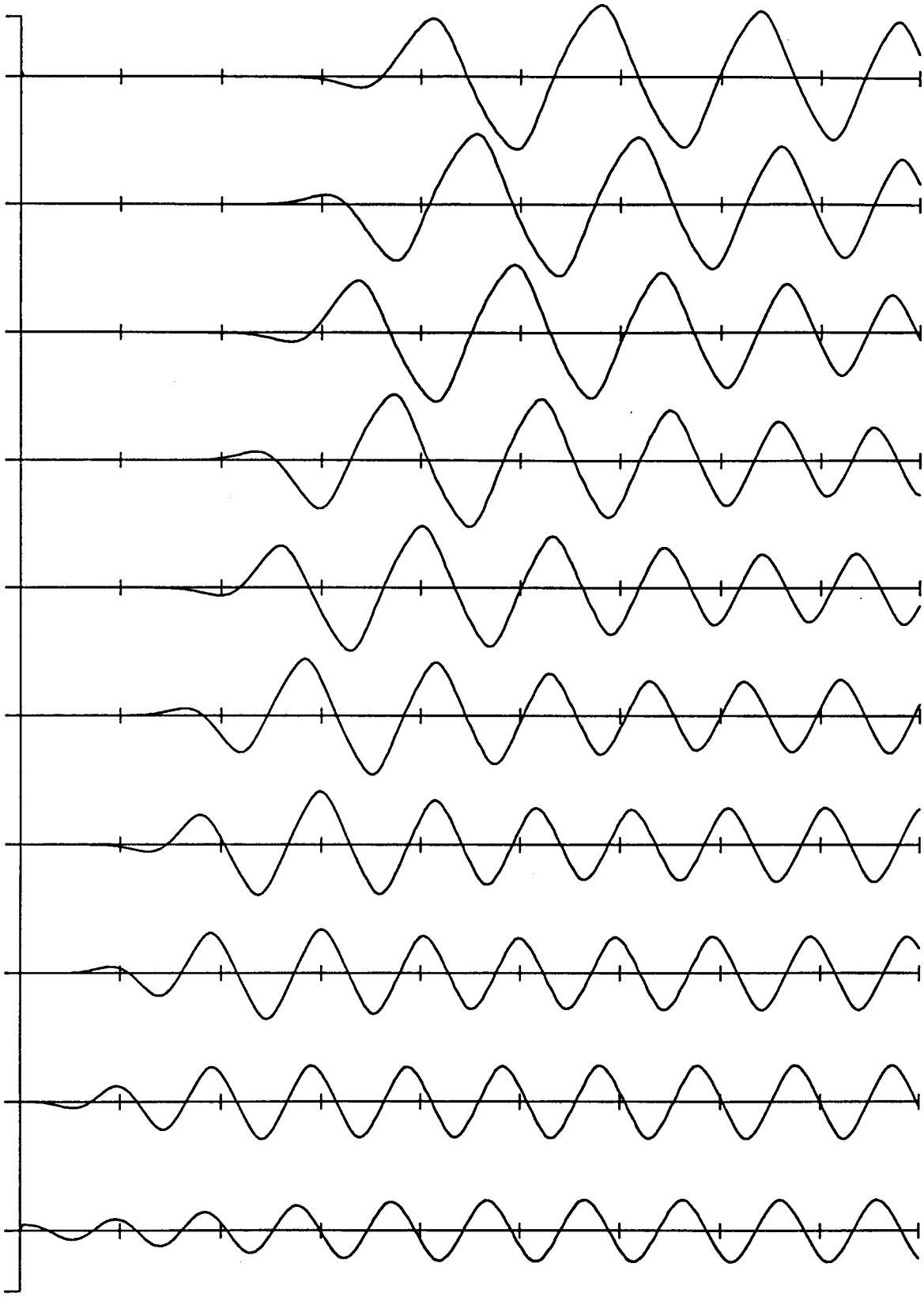

Figure 5.16 - Differential muscle lengths for coupled sections

the qualities observed in the motor nervous system of the organism. This was achieved by using the theoretical framework developed in previous sections and applying the biological constraints imposed by the physical structure of the organism. Some of the features of the current model were dictated by the incompleteness of available biological data, and the success of the model is largely due to the guide provided by the properties of the stochastic model in filling in the gaps. In particular, the statistical underpinnings of the sigmoid transfer function permitted the analysis to concentrate on the significant features of the neural signals while suppressing details that are irrelevant to the overall operation of the system.

The fact that this approach was successful in spite of the incompleteness of the neurophysiological data required for a detailed stochastic model is largely due to the fact that the constraints imposed by the physical properties of the environment, such as the structure of the somatic musculature and the physics of serpentine locomotion, are particularly strong. The structure and function of the motor nervous system are more or less inevitable given these constraints and the nature of the input from the nerve ring.

It is of interest to note that the pattern of connectivity in the nerve cords is minimal. If any of the links shown in figure 5.8 are deleted, the system fails to work. In the actual structure of the nerve cords, shown in figure 5.2, sections of the form shown in figure 5.8

interpenetrate significantly. The net effect is to introduce redundancy and fault tolerance into a system which is structurally minimal. This seems to be a general property of biological nervous systems, and the application of this principle in the above analysis was instrumental in transforming a very difficult problem into a tractable one.

SUMMARY

This thesis presents an application of stochastic techniques to the nervous system of a real organism, with the aim of understanding a simple system in its entirety. This goal has only partially been realized. While the motor postprocessing system was modelled with sufficient specificity to permit its qualitative validation in the light of available electrophysiological data, the sensory preprocessing system and central nervous system models must still be considered speculative. For this reason, this section summarizes the present work as it relates to the motor postprocessing system and the treatment of the other two systems is deferred to the next section.

The Boltzmann machine, with appropriate generalizations, was presented as an ideal theoretical framework for the task. While the construction of a formal model based on the stochastic dynamics of the Boltzmann machine is relatively straight forward, certain practical difficulties presented themselves. Firstly, the biological details are still imperfectly known. As a result, choosing parameter values that result in a quantitatively accurate model proved to be impossible. Some parts of the nervous system of *C. elegans* have been incompletely mapped, necessitating extrapolation of the available biological data. Secondly, while the pattern of synaptic connectivity is known, there is little information upon which to base an assignment of relative synaptic

strength. Thirdly, the implementation of a fully stochastic model is extremely demanding in terms of computational resources and this in itself necessitated the introduction of analytical techniques which neglect some of the detailed stochastic dynamics of the model.

These factors, taken together, prevented the construction of a detailed and quantitatively accurate stochastic model, particularly for the most interesting subsystem, the nerve ring. By utilizing a number of simplifying assumptions, however, fairly satisfying results were obtained with a coarse grained model of the motor nervous system. The action of the system is understood in terms of forces resulting from the energy content of the signals and the correlations between them. While these forces and the response of the system are described in terms of the abstract state space of the system, this viewpoint lays the groundwork for introducing forces that have specific correlates in the physical space inhabited by the organism.

The detailed analysis of the motor postprocessing system uses this technique to create a tractable model of the manner in which the organism produces locomotion. In essence, a gating signal from the nerve ring is transformed into a set of synchronized oscillatory signals which drive the somatic musculature so as to produce waves of flexion which travel along the length of the body. When the system is gated by signals from the nerve ring, it undergoes an effective structural change that produces these oscillations. This is

exactly the sort of context sensitive operation that was described in the analysis of the nerve ring. The system either propagates waves anteriorly or propagates waves posteriorly or relaxes to a quiescent state, depending on the context provided by the central nervous system.

The motor postprocessing system was decomposed into ten idealized sections, each composed of seven interacting state variables. Some of these were neural, while some were mechanical. Inter-section coupling, provided by the distal stretch receptive endings of the DA and VA neurons in the case of reverse locomotion and by those of the DB and VB neurons in the case of forward locomotion, was shown to enforce phase constraints between adjacent sections which lead to travelling waves.

Each section was shown to consist of two pairs of coupled oscillators, each made up of three state variables, as follows. For reverse locomotion, the state variables DA, DE and L form one oscillator, while VA, VE and L form the other. In the case of forward locomotion, DA is replaced by DB and VA is replaced by VB. In both cases, these oscillators are cross coupled by DD and VD so as to maintain a 180 degree phase shift between them. At any time, only one pair is active and the same pair is active in all sections.

Intra-section coupling between the state variables was shown to be modulated by the gating signals from the nerve ring. The closed loop gain of each of these oscillators is insufficient to support sustained oscillations in the absence

of excitation from the nerve ring. The nerve ring selects the direction of motion by driving either DA and VA to produce reverse locomotion or DB and VB to produce forward locomotion. That is to say, by driving one or the other pair of neurons, the central nervous system closes the loops within each section which produce the oscillations and simultaneously selects the direction of coupling between the sections.

Both the means and the variances of the gating signals have a crucial part to play in this process. The noise from the nerve ring diffuses through the nerve cords and determines the slopes of the sigmoid transfer functions of the cord neurons. This effectively sets the gain of these neurons. At the same time, the mean values of these signals diffuse through the nerve cords and set the operating points of the cord neurons.

In the absence of drive from the nerve ring, the cord neurons have relatively narrow sigmoids and the operating point of each one is relatively far to the left of the region in which the sigmoid differs significantly from zero. Thus, the outputs of the neurons remain essentially zero over a wide range values of the input signals from the stretch receptors. In other words, the stretch signals are strongly attenuated and the response of the nerve cords to any variation in these signals is simply to relax back towards a quiescent state.

When a noisy signal from the nerve ring is present, on the other hand, the cord neurons have wider sigmoids and the operating points are in the steep region of the sigmoids. If

the sigmoids are steep enough, the system amplifies the signals coming from the stretch receptors. The feedback loops within each section then give rise to oscillations which grow in magnitude until the signal excursion is such that the peaks drive the neurons into the flatter regions of the sigmoids. This decreases the effective gain of the neurons until sustained oscillations of constant frequency and amplitude result.

If the noise on the gating signals is insufficient, then the sigmoids will be too sharp, and the oscillations of the nerve cord will take on a square waveform. This would require sudden, convulsive contractions of the muscles and result in decreased locomotive efficiency. If there is too much noise on the gating signals, then the sigmoids will be too flat to allow sustained oscillations. Thus, the operation of the motor postprocessing system is seen to depend critically on the mechanism of noise determined variable coupling. Of all neural network architectures, the non-equilibrium Boltzmann machine is the only one which provides this mechanism in a manner which depends only on the stochastic dynamics of the signals being processed.

Discussion

This section discusses some of the more speculative aspects of the techniques presented in this thesis. In particular, the appendices which discuss the sensory preprocessing system and the central nervous system will be summarized below. As far as these two systems are concerned, this thesis has accomplished little more than formulating tractable paradigms and indicating promising directions for further study. The success of these techniques in the case of the motor postprocessing system is heartening, even though it proves little or nothing regarding higher levels of processing. Nonetheless, some of the possible implications of the underlying paradigm with respect to higher levels of processing will be considered.

The sensory preprocessing system was discussed in terms of Kohonen's novelty filter, whereby the system responds most strongly to inputs which are qualitatively new. The behaviour of the organism, the structure of its sensory preprocessing system, and the neurophysiology of its operation were shown to be highly suggestive of this paradigm. As well, some of the theoretical difficulties with this model were shown to be ameliorated by the stochastic characteristics of the Boltzmann machine substrate. The operation of the novelty filter presupposes that the system can be described in terms of projection operators. In a deterministic system, this assumption can not be validated, and the dynamics of the

matrix Riccati equation governing the evolution of the coupling matrix are extremely difficult to characterize. By viewing the system in terms of perturbations from a stochastic background, the conditions under which this assumption holds were clarified.

This viewpoint was carried forward into the analysis of the nerve ring. In addition, the Boltzmann machine model was generalized to allow a different temperature to exist at each of the neurons, and the temperature is characterized in terms of the stochastic properties of the signal. This removes the restrictions imposed by the assumptions underlying equilibrium thermodynamics. Self organisation occurs only when a system is far from equilibrium.

The nerve ring was presented in terms of Sejnowski's skeleton filter, whereby a number of linear systems are embedded in a single nonlinear system. In this paradigm, self organisation takes the form of autonomous transitions between effectively linear systems, giving rise to context sensitive feature association. In the regions between the transitions, the detailed stochastic dynamics of the system can be neglected and the analysis proceeds based on the statistical properties of the signals. The significant parameters are the means and the variations about the means, with the variances determining the degree of significance of the variations.

By symbolically solving the equations of motion of the system, the analysis indicates some of the factors which bear upon this issue and prepares the way for the detailed analysis

of the motor postprocessing system. Although the operation of the motor postprocessing system is a relatively straight forward example of context sensitive processing, it nonetheless illustrates all the salient features, and there seems to be nothing theoretically preventing this approach from being used for the analysis of arbitrarily complex structures whose operation is arbitrarily far removed from the sensory-motor level.

The most interesting regions of state space are those in which the structural transitions take place. That is, those regions in which the border matrix becomes significant and the interaction matrix exhibits a relatively sudden switch from one dynamical regime to another. Unfortunately, these regions are precisely the ones which depend strongly on the higher order statistics of the signals. The present analysis made no attempt to investigate the behaviour of the system in these regions of state space. It merely drew attention to them and showed how they could be recognized. A detailed study of these regions awaits the availability of more computing power and requires the application of more advanced mathematical techniques. These regions of state space provide the most fertile ground for future research.

The behaviour of the system in the transition regions was not characterized theoretically, but it was investigated by simulation. The differences between the fully nonlinear system and its linearized approximation were also investigated by simulation. The results of these simulations bear out the

theoretical analysis. Of particular interest to the biologist are the detailed differences between the two simulations. For example, the differences in asymptotic mean values of the neuron signals and the distortions of the waveforms introduced by the nonlinearities of the system. For the computing scientist, however, the paradigms of the novelty filter and the skeleton filter are of greater interest.

The non-equilibrium Boltzmann machine is derived from the usual model by relaxing the requirement that the link weights be symmetric and introducing the equivalent noise temperatures of the signals in place of a global temperature. This permits the system to remain arbitrarily far from thermodynamic equilibrium and gives rise, therefore, to a dynamical regime in which self organisation can occur. These ideas form a rich theoretical framework which provides useful techniques for the analysis of stochastic information processing.

In this framework, information processing is seen as a process of energy storage and dissipation. Correlated signals give rise to energy storage while uncorrelated signals give rise to energy dissipation. Thus, both the signals and the noise on the signals are important factors in determining the dynamics of the system. Both, therefore, are integral to the processing of information in such a system. The noise level at the input of each neuron chooses a particular sigmoid from the infinite set of transfer functions that the neuron can exhibit, and the mean signal level sets the operating point of the neuron on this curve. Some degree of uncertainty is thus

an inherent property of every item of information.

This approach allows one to characterize the time response of the system to variations in its inputs. Thus, one is able to determine not only what decisions and associations will be made, but how fast and in what order. This is of extreme importance in encoding procedural information in a neural network, in addition to declarative information. It also provides a principled way for procedural and declarative information to interact, in that the modulation of effective link weights by the noise on the signals allows for a smooth blending between control and data. This is based on the idea that the structure of the network is equivalent to control while the signals play the role of data.

It may well be possible, within this framework, to formulate a theory of information processing which includes axiomatic symbol manipulation as a low temperature limiting case. While the behaviour of a system in the regions where the border matrix is negligible closely approximates that of a state machine executing an algorithm, it is highly unlikely that the behaviour of the system can be described in this manner when it is in a transition region. It seems that a close examination of the stochastic dynamics of these regions of state space may permit a rigorous formulation of such presently vague ideas as a smooth transition between algorithms, or superposition of algorithms.

Perhaps the most significant ideas presented in this thesis are the interpretation of information processing in

terms of forces exerted by correlated stochastic signals, decisions implemented by energy dissipation and the manner in which sigmoidal nonlinearities lead to context sensitive processing. This thesis has shown that an information processing paradigm layered on these ideas leads to reasonable explanations of the operation of a real nervous system and has indicated how the study of biological nervous systems can do for computing science what the study of birds' wings has done for aeronautical engineering. The central theme has been the idea that general information processing paradigms can be abstracted from the biological implementation details and provide a guide in the design and construction of stochastic processors that will eventually surpass those which have evolved naturally.

Appendix A - Sensory Preprocessing: Orthogonalization

It is remarkable that a full third of the neurons in the L1 larva are devoted to the anterior sensory apparatus [Ward, Thomson, White & Brenner, 75]. A total of 66 neurons send specialized processes into the sensilla of the head. A few of these also have neuromuscular junctions in the nerve ring and so are members of both sets I and O; the decision to include some such neurons in this section and others in the discussion of motor postprocessing is based on the relative prominence of their sensory endings. By contrast, the entire remainder of the somatic nervous system contains only 22 neurons of a primarily sensory nature, although many of the motoneurons have what appear to be stretch receptive endings [White, Southgate, Thomson & Brenner, 85].

C. elegans is attracted to cyclic nucleotides, anions, cations and hydroxyl ions. It avoids acids, aromatics and carbonate ions [Ward, 76]. The organism responds readily to concentration gradients which the are small enough that the differential concentrations measured laterally between the amphids and/or longitudinally between the amphids and the phasmids would be of insufficient magnitude to drive the processing. The same is true of the temperature gradients detectable by the animal [Ward, Thomson & White, 75]. Having ruled out tropotaxis, Ward suggests a klinotactic mechanism with measurements taken between the extremal displacements of the amphids during the worm's serpentine motion, and assumes

the animal must correlate the position of its head with the signal from its amphids. Indeed, based on the detailed pattern of interconnections between the amphidial neurons, it seems highly unlikely that information regarding the relative strengths of signals at the two amphids is retained by the nervous system.

Unfortunately, the lateral displacements of the amphids during locomotion are insufficient to account for the taxes. Ward proposes this mechanism based on the assumption that a klinokinetic one, wherein the animal compares concentrations at intervals along its path, should be more accurate in faster moving specimens and on the fact that he was able to show that this is not the case. However, it is certainly the ratio between the frequency of turning and the rate of forward motion that is the operative quantity in determining accuracy, and this ratio seems to be more or less constant from one animal to the next under similar conditions, independent of its speed.

An examination of tracks left by *C. elegans* in agar reveals several striking features. Firstly, the paths are straighter in regions of low attractant concentration. Secondly, they are not smoothly curved, but made up of short straight sections separated by comparatively sharp changes in direction. Thirdly, the individual changes in direction display a random character. Fourthly, in the absence of a gradient, each worm seems to have a preferred direction of turning, resulting in looped paths. Finally, the effect of a

concentration gradient seems to be to create an additional bias in the distribution of direction changes.

Together, these observations suggest a klinokinetic orientation mechanism, which operates as follows. In the absence of concentration gradients, the worm proceeds in a relatively straight line, with occasional random direction changes, which may be biased in favour of one direction or the other. This bias need not be constant; it is sufficient for it to be slowly varying. In the presence of a gradient, an additional bias is introduced. While each direction change is still random, the correlation between successive direction changes is determined by the gradient. An increasingly favourable concentration results in a bias towards turning in the same direction as before, while a less favourable one results in the opposite bias. In addition, a more favourable concentration results in an increase in the ratio of turning frequency to rate of locomotion and vice versa.

This scheme places much less stringent requirements on the sensitivity of the amphidial neurons, because the signal signifying a change in concentration builds up as the worm moves forward. When it becomes significant, a direction change is precipitated. This accounts for the increase in ratio of turning rate to locomotion rate in regions of high concentration, and automatically adjusts the sampling rate to suite the available signal strength. Habituation of the sensory system is all that is required to cause this type of behaviour, but the precise mechanism of habituation must be

examined in the light of the complexity of the input signal driving it.

The difficulty concerns the mechanism by which the signal corresponding to the concentration of interest is isolated from the background. No sensory neuron is entirely unimodal. A neuron may be primarily sensitive to a particular substance in the environment, but it will nonetheless respond somewhat to other substances, and such other factors as temperature and mechanical deformation of its ending will cause some response as well. The output signal of the neuron will thus be a complex, and in general nonlinear, superposition of a number of input signals. As a result, the mechanism of habituation must be quite a bit more sophisticated than those which immediately come to mind.

The environmental input signals must be decoupled, and the habituation must be to the decoupled signals. A model exists [Kohonen, 84] which can be adapted to the situation under consideration, consistent with both the structures found in the pattern of interconnections observed in the sensory neurons of *C. elegans* and the stochastic processing paradigm presented above. Before proceeding to discuss the mathematics of the model, it is helpful to consider these structures.

The structure of the anterior sensory apparatus is quite suggestive in terms of functionality. Examination of this structure shows that the pattern of interconnections is fundamentally symmetrical, with random asymmetries. It is interesting to note that the asymmetric connections are denser

on the right than on the left. This could well be a result of the fact that the specimens were cultured on agar, an environment in which they typically lie on their sides. The stimuli from the side in contact with the agar should thus be richer and more complex, resulting in a lateral bias in the complexity of the neural apparatus. One is tempted to conclude that the specimens examined by White et al spent most of their lives lying on their right sides.

Another salient feature of the sensory processing system is the relative sparsity of interconnections between the amphidial neurons and the remainder of the anterior sensory neurons. This argues strongly against an ability to correlate signals from the amphids with signals from the anterior proprioreceptors. The fact that most of amphidial neurons are coupled via gap junctions to their contralateral equivalents forces the conclusion that differential signals between the two amphids are not available to the neuropil in the nerve ring. These observations have strong implications in terms of the possible modes of operation of the system.

On the other hand, there is a very strong coupling between the labial sensilla and the stretch receptors, arguing in favour of a correlation between taste input and head position. Indeed, such a correlation would seem to be necessary to facilitate the ingestion of a bacterium. The IL1 and IL2 neurons send processes to the inner labia which are quite similar, with the exception that the processes of IL2 have access to the chemical milieu immediately inside the

buccal cavity, while those of IL1 do not. It is also significant that the IL2 neurons are primarily presynaptic among the labial neurons, while the others are cross coupled. The interpretation of this structure that immediately suggests itself is that the correlations between inner and outer labial sensilla allow the system to determine whether a touch originates inside or outside the animal's mouth, and because the IL1 neurons form neuromuscular junctions with the muscles of the head, that these neurons form a reflex arc mediating a withdrawal response to touch stimuli at the anterior tip of the worm. The IL2 neurons must then serve to inhibit this reflex when the touch is due to contact with an object that "tastes good".

An examination of the interconnections brings three major features to light. Firstly, the interconnections among the labial neurons and those among the remaining anterior sensory neurons are relatively sparse as compared to the interconnections between these groups. Secondly, the cephalic sensory neurons (CEP) are primarily presynaptic among both groups of neurons, projecting strongly to the inner labial neurons. Thirdly, there is a marked dorso-ventral asymmetry, with denser connections among the dorsal neurons. The distribution of interconnections indicates the importance of correlating information from the anterior stretch receptors with both the inner labial chemoreceptors and the anterior touch receptors. The neurons innervating the cephalic sensilla provide early warning of impending contact with an

obstacle in those nematodes possessing cephalic setae. In *C. elegans*, these sensilla are papillae, but the CEP neurons are nonetheless connected to the other anterior sensory neurons in a manner that seems consistent with an early warning function, perhaps precharging those neurons at which a contact is anticipated. The dorsoventral asymmetry of the anterior sensory apparatus, together with the fact that the organism lies on its side under laboratory conditions, could account for the bias observed in turning behaviour.

Modelling the Sensory Preprocessing System

The lack of information regarding the polarity of the synapses is a serious drawback for a mathematical analysis of the system. A qualitative model can be constructed, however, which can become quantitative when such data become available. Each synapse corresponds to one of the links in the Boltzmann machine model. The exact value of the links is noncritical, because the behaviour of the system remains substantially unchanged when these values are replaced by their signs. The only effect of such a replacement is an increase in overall noise by a factor of $\sqrt{\frac{2}{\pi}}$ [Hopfield, 82]. Thus, a model based on a knowledge of only the positions of synapses and whether each one is excitatory or inhibitory should exhibit the same behaviour as the actual system, with the exception that the model's behaviour would be slightly less certain. The present analysis prepares the way for such knowledge to be incorporated as it becomes available in the form of results of

electrophysiological experiments and immunoassay work.

Not all neurons produce spike trains, and the Boltzmann machine model is expressed in terms of such trains. Hence, it is important to realize that the instantaneous firing rate of a neuron producing a spike train is analogous to the instantaneous membrane potential of one which is not. The expected values of these two variables obey exactly the same dynamics [Sejnowski, 76]. This is particularly important in the nervous system of *C. elegans*, with its predominance of gap junctions with their capability of directly transmitting graded potentials. The present analysis therefore makes no distinction between these two variables, referring to either simply as the activity or output signal of a neuron.

The goal is to determine the vector of signals, $z = \zeta_i$, received by the nerve ring as a result of preprocessing of the input vector, $x = \xi_i$, by the neurons of the anterior sensory apparatus of the organism, taking explicit account of the differential equation of state obeyed by the coupling matrix, $M = \mu_{ij}$, formed by the interconnections between the sensory neurons. When the Boltzmann machine is not at equilibrium, the expected values of the ζ_i will vary with time and, from this point forward, the symbols ζ_i will refer to these variable expectations, or moving averages.

For the sake of simplicity, the effect on the environmental input vectors of the transducer gains of the sensory nerve endings is assumed to be incorporated into the input vector, so that x is the response of the sensory neurons

to the environmental inputs in the absence of any feedback introduced by the interconnections. Both x and z are 66 dimensional vectors, and M is a 66 x 66 matrix.

The analysis is adapted from Kohonen [Kohonen, 84], and will be carried out in two steps. At first, the effect of simple synaptic habituation is considered and shown to lead to orthogonalization of the input signals, whereby the input to the nerve ring becomes the projection of the sensory input onto the orthogonal complement of the space spanned by previous inputs. That is to say, the nerve ring only sees those components of the environmental input which are new. The second step takes into account a forgetting mechanism, which limits the effect of long past inputs.

Before proceeding with this analysis, however, it is necessary to take some steps to bridge the gap between the approach to modification of link weights which is usually taken in the literature and the approach which is adopted here. With this in mind, let the environment be modelled by a set of conditional probabilities. That is, let $P_c(O_n | I_p)$ be the probability that the correct optimal response to input vector I_p is the output vector O_n. On the other hand, let $P_a(O_n | I_p)$ be the probability that the actual response of the system to input I_p will be O_n. It is clear that the desired state of affairs is for these two probabilities to be equal for all n and p, so the distance between the actual behaviour of the system and the optimal one is the distance between the conditional distributions P_a and P_c. Information theory

[Kullback, 68] provides a distance measure on the space of probability distributions, called the asymmetric divergence. Applied to the current problem, this measure is expressed by

$$G = \sum_{n,p} P_c(O_n \wedge I_p) \ln\left(\frac{P_c(O_n|I_p)}{P_a(O_n|I_p)}\right) \qquad (A.1)$$

and is easily seen to be nothing other than the difference in information content between the two distributions. It is not a metric because its value depends on which distribution is taken to be prior and which is taken to be posterior. In the present context, the structure of the environment implicitly supplies the *a priori* distribution and G measures the accuracy of the *a posteriori* estimate of it embodied in the behaviour of the system. It should be noted that the *a priori* distribution represents not the environment itself, but a model of how to deal effectively with the environment, and the *a posteriori* distribution represents what the system has learned about this model.

Now, the only way for the nodes in I to affect those in O is through the W_{ij}. Hence, in order to minimize G, it is sufficient to find the partials of G with respect to the W_{ij} and set them to zero. Since the P_c are independent of the internals of the system,

$$\frac{\partial G}{\partial W_{ij}} = -\sum_{n,p} \frac{P_c(O_n \wedge I_p)}{P_a(O_n \wedge I_p)} \frac{\partial}{\partial W_{ij}} P_a(O_n \wedge I_p) \qquad (A.2)$$

which is the sum of the fractional rates of change of the actual conditional probabilities, weighted by the correct conditional probabilities. From this expression, it can be shown that

$$\frac{\partial G}{\partial W_{ij}} = -\frac{1}{kT}\left(\sum_{mnp} \zeta_i^{mnp}\zeta_j^{mnp} P_c(H_m \wedge O_n \wedge I_p) - \sum_{rst} \zeta_i^{rst}\zeta_j^{rst} P_a(H_r \wedge O_s \wedge I_t)\right) \quad (A.3)$$

but these sums represent the probabilities that nodes i and j are simultaneously active in the actual operation of the system and during optimal operation, respectively. That is to say, these sums represent correlations between node activities. Denoting these probabilities by p^c_{ij} and p^a_{ij}, equation (A.3) reduces to

$$\frac{\partial G}{\partial W_{ij}} = \frac{1}{kT}(p^a_{ij} - p^c_{ij}) \quad (A.4)$$

This is an extremely significant relation, because it shows that the global behaviour of the system can be adjusted in a meaningful and principled way by changing the W_{ij} according to purely local information. Aside from proving that a Boltzmann machine realizes a Markov random field [Isham, 81] [Geman & Geman, 84], equation (A.4) also suggests a method by which the system can learn the higher order constraints in its environment.

By periodically adjusting the W_{ij} based on the distance between the actual and desired probabilities, moving in a direction opposite to the gradient of G since its minimum is

sought,

$$\Delta W_{ij} = \varepsilon (p^{c}_{ij} - p^{a}_{ij}) \qquad (A.5)$$

and thus it is clear that the partials will eventually vanish. There is only one problem with this scheme. It requires a prior knowledge of the desired global behaviour, even if only on the part of a "teacher" to periodically demonstrate it. In this sense, the claim that all information is purely local is misleading, because the teacher provides global information. In order to implement the scheme expressed by equation (A.5), each node must be able to determine, at every instant, whether it is sampling from p^c or p^a.

An alternate approach, which avoids the need for a teacher, is obtained by observing that an unbiased estimate of the above probabilities can be found simply by observing the states of nodes i and j for relatively short intervals separated by comparatively longer periods. All that is required is one additional piece of information: the mean effectiveness of the current behaviour of the system over each interval, which can be modelled as a number between $-\varepsilon$ and ε. This information is the only essential thing provided by the teacher. Then, denoting the effectiveness by α, if

$$\Delta W_{ij}(t) = \alpha(t) \langle \zeta_i(t) \zeta_j(t) \rangle \qquad (A.6)$$

is added to W_{ij} at each firing of node i, then the long term average effect will be equivalent to that of equation (A.5). In this expression, t is the midpoint of the interval, and the

averaging indicated on the right hand side is carried out over the length of the interval. This amounts to substituting a stochastic gradient approach for a strict gradient descent [Honig & Messerschmitt, 84], and can be applied even when the environmental probabilities do not have stationary statistics.

In terms of a nervous system, the effectiveness can be easily defined. It is the degree to which a response of the system succeeds in bringing about a result which is beneficial to the organism. If the organism finds itself in a concentration gradient of some noxious chemical, behaviour that results in movement up the gradient has negative effectiveness and vice versa. Of course, several problems may need solving at once. The organism may be hungry, tired and cold, for example. The effectiveness of behaviour with respect to each of these problems is encoded by a different subset of the endocrine system which, by modulating the activity level of the nervous system, controls at least the rate of synaptic plasticity.

It can be seen from the above discussion that learning of arbitrary high order constraints in the environment can be achieved in a principled way by modifying local behaviour in response to global effectiveness. The modifications are performed in a uniform way, completely independent of the problem domain, driven solely by the information which needs to be processed, without any *a priori* knowledge of what that information may be. The only limitations are imposed by the suite of sensory modalities and actuator capabilities. Thus

a Boltzmann machine is an ideal model for a nervous system and provides precisely the sort of substrate for self organization that was sought in the introduction.

In the case of the sensory preprocessing system, the high order structure in the input signals which the system needs to adapt to are the nonlinear superpositions of sensory signals discussed above. The optimal response of the sensory preprocessing system is defined by the input requirements of the central processing system. This in turn, may be characterized by the requirement not to be swamped by a stream of irrelevant input signals while being notified immediately of significant sensory events. These constraints define the optimal dynamics of the sensory preprocessing system.

Clearly, the variation of the output of the sensory preprocessing system is most strongly dependent on the neurons operating near their thresholds, and for these neurons, the output can be written as

$$\zeta_i = \xi_i + \sum_j \mu_{ij}\zeta_j \tag{A.7}$$

and, if synapses decay simply by depletion of the supply of neurotransmitters, the coupling matrix satisfies

$$\frac{d\mu_{ij}}{dt} = -\alpha\zeta_i\zeta_j \tag{A.8}$$

where α is a rate constant which can be determined by simple measurements on individual synapses. It is revealing to

compare the form of this expression to that of equation (A.6) in order to see just how broad a range of interpretations can be placed on the effectiveness parameter. In a system whose purpose is to give no output for familiar input signals, any output can be thought of as indicating negative effectiveness. The expected response of the system is then to learn a function of the new input signal so that the output may return to zero. Expressed in matrix notation, the above equations assume the form

$$z = x + Mz \tag{A.9}$$

and

$$\frac{dM}{dt} = -\alpha z z^T \tag{A.10}$$

Due to the sigmoid transfer function given by equation (3), the components of z are all nonnegative. Hence each of the elements of the right hand side of equation (A.10) is negative until z = 0 and then adaptation ceases. As a result, those neurons to which equation (A.9) does not initially apply are gradually brought into the region of the sigmoid in which it does. Eventually, all of the sensory neurons are brought to a state in which their output is essentially zero. At this point, we have

$$z = (I-M)^{-1}x = 0 \tag{A.11}$$

where I is the identity matrix and $P = (I-M)^{-1}$ is called the

transfer operator of the sensory preprocessing system. It is related to the interaction matrix, which is discussed in the next section, in the context of the nerve ring.

Essentially, equation (A.11) expresses that the input, x, is projected onto the orthogonal complement of the subspace spanned by x. Thus, P is a projection matrix. Now, suppose that x changes from x_0 to x_1, such that

$$x_1 = ax_0 + y_0 \tag{A.12}$$

where a is some constant and y_0 lies in the orthogonal complement of the subspace spanned by x_0. Then

$$z_1 = aPx_0 + Py_0 \tag{A.13}$$

but the first term on the right hand side of this equation is zero by equation (A.11), hence

$$z_1 = Py_0 \tag{A.14}$$

and the output of the sensory preprocessing system is now determined by that component of the input which is new, in the sense that it is not simply a multiple of the old input. That is to say, the sensory preprocessing system responds to that component of the input signal which is qualitatively new with respect to past inputs.

This analysis assumes that the input to the sensory preprocessing system is piecewise constant. In fact, the input will be slowly varying with occasional rapid changes. As a result, the transfer operator, P, will never exactly be

a projection matrix, although it continuously relaxes towards one.

Nonetheless, the approximation will be valid as long as the input varies slowly enough to ensure that the output of the sensory preprocessing system is small compared to the noise in the system. This noise is composed of random fluctuations in the input signals as well as random interactions between the neurons themselves. It is important to note that it is precisely when the input varies in such a manner as to cause the approximation to fail for a short period that the sensory preprocessing system reports to the nerve ring that a significant sensory event has taken place. In other words, as long as the input varies slowly enough that the adaptation of the sensory preprocessing system can keep up with it, the transfer operator will approximate a projection operator and the nerve ring will receive no significant input.

In order for the system to continue operating for any appreciable length of time, the supply of neurotransmitters at the synapses must be replenished. Otherwise there will come a time when no input signal can drive the output away from zero. This amounts to an infinitely long memory, which will eventually cause the sensory preprocessing system to reach a state in which it has seen all possible inputs; a state, therefore, in which nothing is new. The replenishment of the supply of neurotransmitters thus acts as a forgetting mechanism.

If this occurs at a constant rate, independent of the

activity of the system, then

$$\frac{d\mu_{ij}}{dt} = \beta - \alpha \zeta_i \zeta_j \tag{A.15}$$

and the system will reach a steady state when

$$\zeta_i \zeta_j = \frac{\beta}{\alpha} \tag{A.16}$$

In order for the preceding analysis to remain valid under these conditions, the correlations of the residual output signals must be effectively buried in the noise. In other words, the correlations of the residual signals must be significantly less than the correlations of the random fluctuations in the outputs of the neurons about their expected values. That is,

$$\langle (\zeta_i - \langle \zeta_i \rangle)(\zeta_j - \langle \zeta_j \rangle) \rangle \gg \frac{\beta}{\alpha} \tag{A.17}$$

This condition can only be verified by further biological experiment. Nonetheless, it represents an intuitively appealing paradigm for the operation of the sensory preprocessing system, based on biologically reasonable assumptions.

The view which emerges from this analysis is of a set of small, essentially uncorrelated signals presented to the nerve ring by the sensory preprocessing system under steady state conditions. A significant sensory event results in a flurry

of correlated activity, given by equation (A.14), which is damped as described by equation (A.15), causing it to decay back to the essentially uncorrelated condition described by equation (A.17). In this view, the nerve ring is informed of a series of more or less distinct sensory events by perturbations in a relatively stationary stochastic background.

Appendix B - The Nerve Ring: Context Sensitive Association

The neuropil of the nerve ring is located in the cervical region of the organism, between the two bulbs of the pharynx. It is more or less contiguous with the retrovesicular ganglion, which in turn gives rise to the ventral and dorsal nerve cords. It performs the major integrative processing of the nervous system. Many of the neurons which are classified as ring interneurons send processes through the retrovesicular ganglion and into the nerve cords, exhibiting significant synaptic connectivity to one another throughout. For purposes of analysis, all of these neurons are considered to be part of the nerve ring.

The processes of the ring interneurons are situated circumferentially around the pharynx, and form synapses and gap junctions to one another in passing. There is typically no presynaptic thickening of the neuronal processes. The processes typically have an angular extent of about 90 degrees, but this extent varies between about 60 degrees and about 270 degrees.

Most of the neurons in the nervous system of *C. elegans* are located in or send processes into the nerve ring. Discounting those which make up the sensory preprocessing system and those which innervate the cervical muscles, there are 69 neurons which can be classified unambiguously as interneurons. These neurons seem to be exclusively cholinergic, so that all synapses between ring interneurons are

excitatory. As seems often to be the case, the connectivity among the neurons in *C. elegans* is on the order of the square root of the number of neurons in the system. That is, each of the 239 neurons in the system receives input from about 15 others.

Modelling the Nerve Ring

The analysis carried out in this section will establish the qualitative dynamics of the nerve ring, making use of the characteristics of the input from the sensory preprocessing system as derived in the previous section. It will be shown that the nonlinear interactions between the neurons lead to an embedding of a number of approximately linear systems in the single nonlinear one. The state space of the system can thus be partitioned into two types of regions: those in which the system can be treated as a linear filter and those in which the system makes a transition from one filter to another.

Since the *modus operandi* of a linear filter can be described in terms of feature association, and a particular linear filter always associates the same output with a given input, such a filter performs context independent feature association. However, since the input can drive the system through a transition from one linear filter to another, with a corresponding change in the association between input and output, the system is capable of varying its associations in a way that depends on certain parameters of the input signal. That is, the system performs context sensitive feature

association. The remainder of this section presents a characterisation of the transition regions and a detailed analysis is presented of the time response of the system outside of the transition regions.

As in the previous section, let the transfer function of each neuron be the sigmoid given by equation (3). Then we may write

$$\eta_i = \xi_i + \sum_j \mu_{ij}\zeta_j \tag{B.1}$$

where

$$\zeta_i = p(\eta_i) \tag{B.2}$$

Equation (A.1) is an approximation to these two equations for neurons operating near threshold and η is equivalent to ΔE of equation (3) and represents the effective membrane potential of the neuron. The temperature from equation (3) can also be given a meaning in terms of the membrane potential. Since equation (3) is just the cumulative Boltzmann distribution, the variance of η can be found by integrating η conditioned by the Boltzmann density function:

$$\langle \eta_i \rangle^2 = \frac{1}{kT_i}\int_{-\infty}^{+\infty} x^2 p(x)\,(1-p(x))\,dx = \frac{\pi^2}{3}(kT_i)^2 \tag{B.3}$$

This represents an extremely significant generalization of the Boltzmann machine model in that, by allowing the temperature to differ from node to node, it opens the way to analysis of

its nonequilibrium behaviour. Equation (B.3) rests on the idea of the equivalent noise temperature of a signal, which reduces in this case to

$$kT_i = \frac{\sqrt{3}}{\pi}\sigma_i \qquad (B.4)$$

where σ is the variance of the membrane potential.

As can easily be verified numerically, the asymmetric divergence between the Boltzmann and Gaussian distributions is minimized when the variance of the Gaussian satisfies equation (B.4). In this case, the difference in information content between the two distributions is less than half of a bit, and the cumulative distributions never differ by more than 1% over their entire range. This will be of importance in the analysis which follows, wherein certain steps depend on the distribution being approximately Gaussian.

In equation (B.1), the ξ represent the inputs to the nerve ring from the sensory preprocessing system. Following Sejnowski [Sejnowski, 76], we assume that these inputs have approximately stationary expectations, in other words that the $<\xi>$ are constant, and consider what occurs when the ξ are perturbed from their equilibrium values. This is consistent with the paradigm presented in the previous section for the operation of the sensory preprocessing system.

Let the inputs depart from equilibrium along a curve in state space parameterized by s, such that the ξ vary smoothly with s and s = 0 at equilibrium. In terms of this parameter,

the state of the system can be expanded in a Taylor series as follows:

$$\eta_i(s) = \eta_i(0) + s\frac{d\eta_i}{ds}(0) + \frac{s^2}{2}\frac{d^2\eta_i}{ds^2}(0) + \ldots \tag{B.5}$$

Here η (0) is the equilibrium membrane potential, $\frac{d\eta_i}{ds}(0)$ is the first order variation of the membrane potential, and so on. Since the μ_j do not vary with s, we have

$$\frac{d\eta_i}{ds} = \frac{d\xi_i}{ds} + \sum_j \mu_{ij}\frac{d\zeta_j}{d\eta_j}\frac{d\eta_j}{ds} \tag{B.6}$$

but it is easily verified that

$$\frac{d\zeta_j}{d\eta_j} = \frac{1}{\sigma_j}p(\eta_j)(1-p(\eta_j)) \tag{B.7}$$

where the factor of $\frac{\sqrt{3}}{\pi}$ has been absorbed into σ.

The form of equation (B.6) suggests that from the coupling matrix, $M = \mu_j$, we derive a new matrix, $N = \nu_j$, defined by

$$\nu_{ij} = \frac{\mu_{ij}}{\sigma_j}p(\eta_j)(1-p(\eta_j)) \tag{B.8}$$

This new matrix is called the interaction matrix and, in terms of it, equation (B.6) reads

$$\frac{d\eta_i}{ds} = \frac{d\xi_i}{ds} + \sum_j \nu_{ij}\frac{d\eta_j}{ds} \tag{B.9}$$

The second order variation is similarly derived:

$$\frac{d^2\eta_i}{ds^2} = \frac{d^2\xi_i}{ds^2} + \sum_j \mu_{ij}\left(\frac{d^2\zeta_j}{d\eta_j^2}\left(\frac{d\eta_j}{ds}\right)^2 + \frac{d\zeta_j}{d\eta_j}\frac{d^2\eta_j}{ds^2}\right) \tag{B.10}$$

But by differentiating equation (B.7),

$$\frac{d^2\zeta_j}{d\eta_j} = \frac{1}{\sigma^2}p(\eta_j)(1-p(\eta_j))(1-2p(\eta_j)) \tag{B.11}$$

and we define $B = \beta_j$ by

$$\beta_{ij} = \frac{\mu_{ij}}{\sigma_j^2}p(\eta_j)(1-p(\eta_j))(1-2p(\eta_j)) \tag{B.12}$$

This matrix is called the border matrix. In terms of this matrix, equation (B.10) reads

$$\frac{d^2\eta_j}{ds^2} = \frac{d^2\xi_i}{ds^2} + \sum_j \beta_{ij}\left(\frac{d\eta_j}{ds}\right)^2 + \sum_j \nu_{ij}\frac{d^2\eta_j}{ds^2} \tag{B.13}$$

In matrix notation, equation (B.9) reads

$$\eta' = \xi' + N\eta' \tag{B.14}$$

or just

$$\eta' = A^{-1}\xi' \tag{B.15}$$

where

$$A^{-1} = (I-N)^{-1} \tag{B.16}$$

is the first order transfer operator of the nerve ring, giving

its response to the first order variation in the input from the sensory preprocessing system. Similarly, equation (B.13) reads

$$\eta'' = \xi'' + B(\eta')^2 + N\eta'' \tag{B.17}$$

in matrix notation. That is,

$$\eta'' = A^{-1}(\xi'' + B(\eta')^2) \tag{B.18}$$

Now, this process can be carried on indefinitely to successively higher orders of variation, and the results substituted back into equation (B.5). The result, as long as B is negligible, is a Taylor series for ξ, multiplied from the left by A^{-}. But this is just equation (A.5) again, except that the slopes of the sigmoids are taken explicitly into account. Thus, when the border matrix is negligible, the response of the system is completely determined by the interaction matrix. The matrices, N and B, contain a great deal of information about the response of the nerve ring to inputs from the sensory preprocessing system.

The elements of the interaction matrix are significant only for those neurons with membrane potentials near threshold. This can be seen by examining figure B.1 b. Each column of N is conditioned by a value chosen from a curve of the form indicated in this figure. The value is completely determined by the fact that the membrane potential selects the horizontal coordinate while its variance selects the width and height of the curve. Only synapses from neurons for which

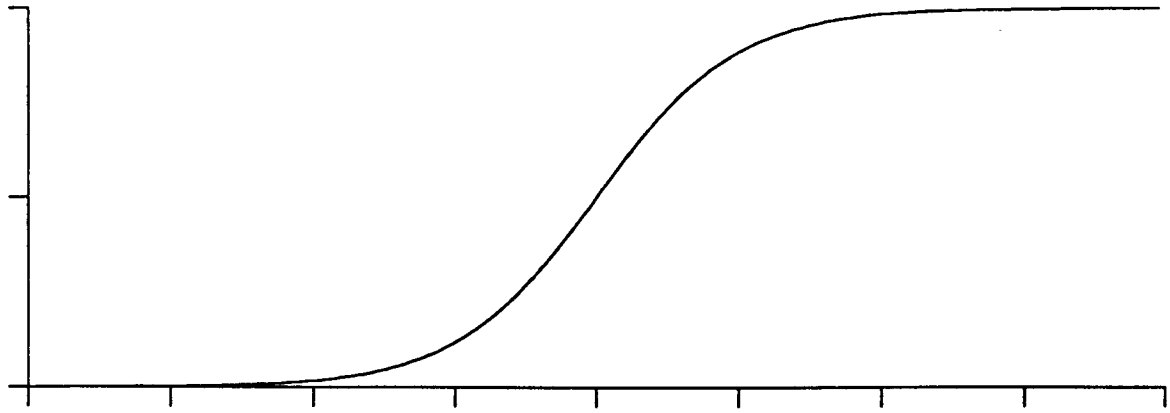

Figure B.1 a - The sigmoid transfer function of a neuron

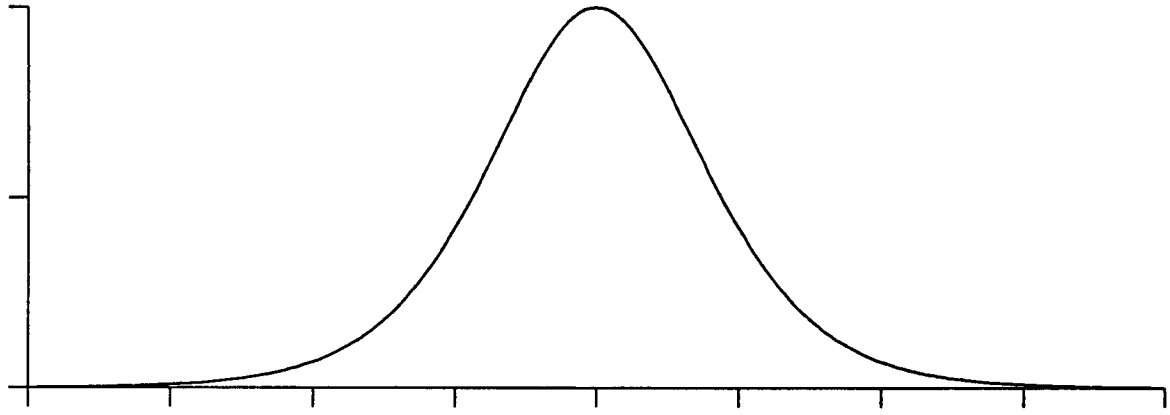

Figure B.1 b - The derivative of the sigmoid

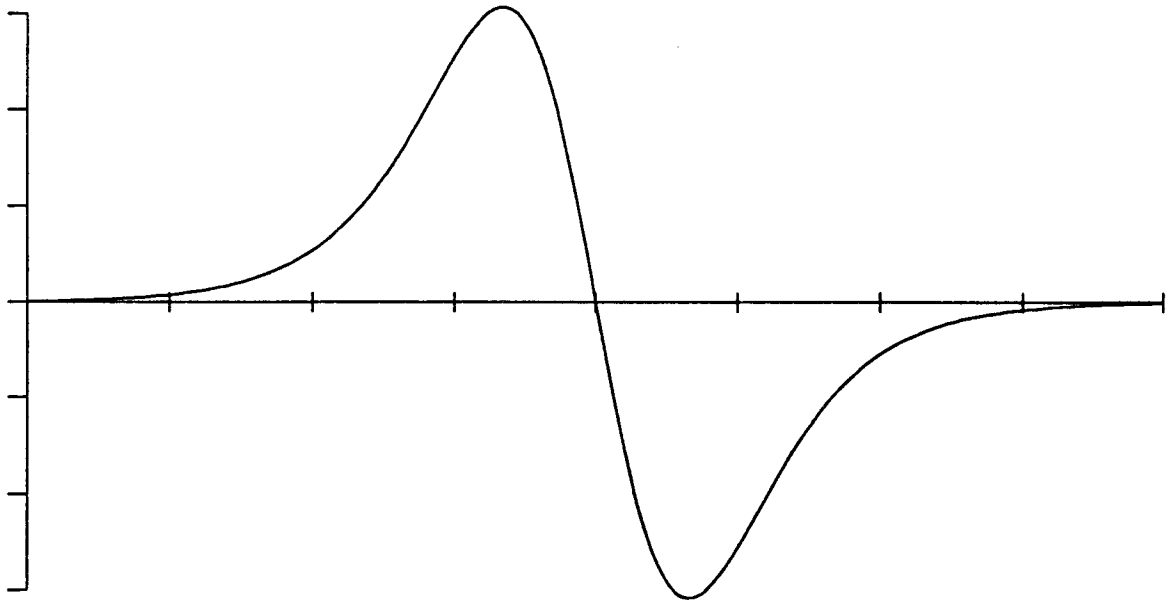

Figure B.1 c - The second derivative of the sigmoid

this value is significantly different from zero contribute significantly to the change in input to a given neuron.

Essentially, the interaction matrix picks out a subset of the neurons in the nerve ring, called critical neurons, which exhibit a significant change in state in response to a change in the input from the sensory preprocessing system. All other neurons exhibit a negligible change in state. Thus the transfer function of the nerve ring is context sensitive, being determined by the constant or slowly varying background activity in the ring interneurons via equation (B.8).

Similarly, the transition from one transfer function to another is determined by the border matrix, B. Referring to figure B.1 c, the significant elements of the border matrix are those associated with neurons for which the curve in this figure differs appreciably from zero. These neurons are called border neurons. Border neurons are biased by the background so as to be easily forced into or out of the set of critical neurons by an input perturbation. Those nearest to the peaks of the curve in figure B.1 c are the ones which contribute most strongly to a transition from one transfer function to another.

As long as the input from the sensory preprocessing system varies in such a way as to cause no migration between the critical and border neurons, the transfer function of the nerve ring remains essentially unchanged. In this case, the dynamics of the system remain qualitatively the same over a range of inputs. Whatever behaviour results from the output

of the nerve ring to the motor postprocessing system, the association between stimuli and responses will be modulated in degree, but will not change in kind. On the other hand, when a migration does occur between the critical and border neurons, the dynamics of the system passes through a bifurcation and the association between stimulus and response will exhibit a sudden change in kind.

Time Response of the Nerve Ring

Within this qualitative dynamical framework, then, we may proceed to analyze the time dependence of the outputs from the nerve ring in response to its inputs. In the foregoing analysis, the tacit assumption was made that all signals vary slowly enough that the time constants of the neurons may be neglected. This amounts, in the thermodynamic framework, to assuming that the changes are reversible. But reversible changes are assumed to always occur slowly enough that the equilibrium is not significantly disturbed. In other words, the distribution always has time to relax back to its equilibrium form. This assumption must be relaxed in order to apprehend the time development of the system state because all time dependence of the system state variables vanishes at equilibrium.

Due to the storage of charge in its membrane capacitance, a neuron integrates its input signal over time. Due to its membrane resistance, this charge leaks away over time. The result is that the neuron acts as a lossy integrator with a

time constant, τ, on the order of 15 msec. The membrane potential of a single, undriven neuron thus satisfies

$$\tau \frac{d\eta}{dt} + \eta(t) = 0 \tag{B.19}$$

which has the solution

$$\eta(t) = \eta_0 e^{-\frac{t}{\tau}} \tag{B.20}$$

where η is the initial membrane potential. That is, in the absence of input, the membrane potential decays exponentially to zero with time constant τ. When the neuron is driven by an input, or forcing function, f(t), equation (B.19) reads

$$\tau \frac{d\eta}{dt} + \eta(t) = f(t) \tag{B.21}$$

Multiplying through by the inverse of the solution to the homogeneous equation, gives

$$e^{\frac{t}{\tau}}\left(\tau \frac{d\eta}{dt} + \eta(t)\right) = e^{\frac{t}{\tau}} f(t) \tag{B.22}$$

but the left hand side of this equation can be rewritten as a direct time derivative, thus:

$$\tau \frac{d}{dt}\left(e^{\frac{t}{\tau}} \eta(t) - \eta_0\right) = e^{\frac{t}{\tau}} f(t) \tag{B.23}$$

which can be integrated immediately to yield

$$e^{\frac{t}{\tau}}\eta(t) = \eta_0 + \frac{1}{\tau}\int_0^t e^{\frac{s}{\tau}} f(s)\,ds \tag{B.24}$$

and therefore

$$\eta(t) = \eta_0 e^{-\frac{t}{\tau}} + \frac{1}{\tau} e^{-\frac{t}{\tau}} \int_0^t e^{\frac{s}{\tau}} f(s)\,ds \tag{B.25}$$

where the first term on the right hand side expresses the decay of the initial membrane potential and the second term gives the response to the forcing function alone.

Now, when the neurons are coupled via the coupling matrix, M, the expected value of the membrane potentials satisfies

$$\tau \frac{d}{dt}\langle \eta_i \rangle + \langle \eta_i \rangle = \langle \xi_i \rangle + \sum_j \mu_{ij} \zeta_j \tag{B.26}$$

which has the same form as equation (B.21), with the sum of the input from the sensory preprocessing system and the feedback from the other neurons acting as the forcing function for each neuron. When the system is at equilibrium the background is stationary, so that

$$\frac{d}{dt}\langle \eta_i \rangle = 0 \tag{B.27}$$

and we recover equation (B.1). However, by an argument similar to that which lead to equation (B.14), the response to a perturbation in the input from the sensory preprocessing

system will satisfy

$$\tau\frac{d}{dt}(\eta-\langle\eta\rangle) + A(\eta-\langle\eta\rangle) = \xi-\langle\xi\rangle \tag{B.28}$$

where the matrix, A, is defined as in equation (B.16). This equation is analogous to equation (B.21), except that scalars have been replaced by vectors. The solution is conceptually similar, except that some complexity is introduced by the fact that A is a matrix and is not generally diagonalizable.

Proceeding as before, we form the homogeneous equation

$$\tau\frac{d}{dt}(\eta-\langle\eta\rangle) + A(\eta-\langle\eta\rangle) = 0 \tag{B.29}$$

Now, since the nerve ring is a 69 dimensional system, there are 69 linearly independent solutions to this equation, from the superposition of which we will construct the solution to equation (B.28). Thus, we introduce an integrating factor of the form

$$T^{-1} = e^{-A\frac{t}{\tau}} \tag{B.30}$$

where the matrix exponential is defined by

$$e^{At} = \sum_{k=0}^{\infty}\frac{t^k}{k!}A^k \tag{B.31}$$

It is easily verified that

$$\frac{d}{dt}e^{At} = Ae^{At} \tag{B.32}$$

and therefore T satisfies a homogeneous matrix differential equation analogous to equation (B.29). That is,

$$\tau \frac{d}{dt} T(t) + AT(t) = 0 \tag{B.33}$$

Therefore, T^- does also and, upon multiplying equation (B.28) through by $T^-(t)$, we obtain an analog of equation (B.23). Namely,

$$\tau \frac{d}{dt} \left(T^{-1}(t) (\eta - \langle \eta \rangle) - (\eta(0) - \langle \eta \rangle) \right) = T^{-1}(t) (\xi - \langle \xi \rangle) \tag{B.34}$$

which we can now integrate to obtain

$$\eta - \langle \eta \rangle = T(t) (\eta(0) - \langle \eta \rangle) + T(t) \int_0^t T^{-1}(s) (\xi(s) - \langle \xi \rangle) \, ds \tag{B.35}$$

but since T(t) does not depend on s, it may be brought into the integrand, and since it is easily verified that

$$T(t) T^{-1}(s) = T(t-s) \tag{B.36}$$

equation (B.35) reduces to

$$\eta(t) - \langle \eta \rangle = T(t) (\eta(0) - \langle \eta \rangle) + \int_0^t T(t-s) (\xi(s) - \langle \xi \rangle) \, ds \tag{B.37}$$

This expression gives the complete response of the system to a vector of input signals, as a function of time. The matrix, T(t), is called the impulse response of the system. It gives the response of the system at time, t, to an input that was

present at t = 0. Similarly, T(t-s) gives the response of the system at time, t, to an input that was present at time, s. Hence the first term on the right hand side of equation (B.37) gives the decaying response of the system to the initial conditions and the second term gives the net effect at time, t, of all the inputs since t = 0.

Characterization of the Impulse Response

The term impulse response means that the matrix, T(t), expresses the response of the system to the input vector whose components are all unit impulses located at t = 0. This matrix completely characterizes the system in those regions of state space in which no migration occurs between critical and border neurons. The columns of T must therefore be linear combinations of the solutions of the homogeneous equation, equation (B.29), with the property that

$$T(0) = I \tag{B.38}$$

In order to characterize these solutions, it is convenient to work in a basis in which A, and therefore T, have the simplest possible form. To this end, we recall that every linear operator can be expressed as the sum of a diagonalizable operator and an operator which is nilpotent on a set of mutually exclusive and jointly exhaustive invariant subspaces. We therefore proceed to find the matrix, S, which transforms A into a form which exhibits this structure explicitly:

$$J = S^{-1}AS = D+C \tag{B.39}$$

where D is diagonal and C is nilpotent on each of the invariant subspaces of A. Since the dynamics of the system do not depend on the coordinate system used to describe them, the change of coordinates expressed by the matrix, S, merely has the effect of simplifying the process of finding a solution to equation of motion of the system, equation (B.28). Since equation (B.39) is a similarity transform, the invariant subspaces of J are the same as those of A.

The matrix J is called the Jordan canonical form of A and has the form

$$J = \begin{bmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_n \end{bmatrix} \tag{B.40}$$

where each of the blocks, J_r is the restriction of J onto the set of invariant subspaces associated with the possibly complex eigenvalue λ of A, where

$$\lambda_r = \alpha_r + i\omega_r \tag{B.41}$$

In turn, each of the J_r has the form

$$J_r = \begin{bmatrix} J_{r1} & 0 & \cdots & 0 \\ 0 & J_{r2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{rm} \end{bmatrix} \tag{B.42}$$

where each of the blocks J_{rm} is the restriction of J onto the single invariant subspace generated by the eigenvector, Φ, of A. This block has the form

$$J_{rs} = \begin{bmatrix} \lambda_r & 1 & 0 & \cdots & 0 \\ 0 & \lambda_r & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & \lambda_r \end{bmatrix} \tag{B.43}$$

Recall that the eigenvalues and eigenvectors of a linear operator are those scalars and vectors which satisfy

$$A\Phi_{rs1} = \lambda_r \Phi_{rs1} \tag{B.44}$$

In other words, each eigenvalue has associated with it one or more eigenvectors which have the property that, when operated on by A, they are simply scaled by the eigenvalue. The column position, u , of Φ in the matrix, S, is the same as the first column of J in J, and Φ generates the invariant subspace of J as follows. Let d be the dimensionality of the subspace. If d = 1 then Φ is the single basis vector of the subspace. Otherwise the remaining d - 1 basis vectors, Φ , occupy successive columns of S and it can be seen that, since AS = SJ,

$$A\Phi_{rsk} = \lambda_r \Phi_{rsk} + \Phi_{rsk-1} \tag{B.45}$$

Equations (B.44) and (B.45) together imply that

$$(A-\lambda_r I)^k \Phi_{rsk} = 0 \tag{B.46}$$

These vectors are called the generalized eigenvectors of A. A dual set of generalized eigenvectors, Ψ , can be constructed to form the rows of S^- . Since $S^- A = JS^-$, these

vectors satisfy

$$\Psi_{rsk}^{T}(A-\lambda_r I)^{k} = 0 \tag{B.47}$$

The vector Ψ^T occupies row position u in S^- and the remaining d - 1 vectors occupy successive rows.

Now, it is easily verified that

$$e^{At} = Se^{Jt}S^{-1} \tag{B.48}$$

that is,

$$e^{At} = Se^{Dt}e^{Ct}S^{-1} \tag{B.49}$$

We therefore proceed to determine the structure of this matrix. For a single Jordan block,

$$J_{rs} = D_{rs} + C_{rs} \tag{B.50}$$

and it is immediate from equation (B.31) that

$$e^{D_{rs}t} = e^{\lambda_r t}\delta_{ij} \tag{B.51}$$

and thus $e^{D_{rs}t}$ is a diagonal matrix whose nonzero elements are $e^{\lambda_r t}$. The situation for C is somewhat more involved. Since

$$C_{rs} = \delta_{ij-1} \tag{B.52}$$

it follows that

$$C_{rs}^{k} = \delta_{ij-k} \tag{B.53}$$

and therefore that

$$C_{rs}^{d_{rs}} = 0 \tag{B.54}$$

hence the series expansion of $e^{C_{rs}t}$ terminates after d terms. That is,

$$e^{C_{rs}t} = \sum_{k=0}^{d_{rs}-1} \frac{t^k}{k!} C^k \tag{B.55}$$

or just

$$e^{C_{rs}t} = \frac{t^{j-i}}{(j-i)!} \sum_{k=1}^{j} \delta_{ik} \tag{B.56}$$

and thus $e^{C_{rs}t}$ is an upper triangular matrix with elements of the form $\frac{t^k}{k!}$ on the k superdiagonal. When this matrix is multiplied by $e^{D_{rs}t}$ all of its elements are simply multiplied by $e^{\lambda_r t}$.

The functions $\frac{t^k}{k!} e^{\lambda_r t}$ are called the eigenfunctions of A and the columns of the solution of equation (B.33) are composed of linear combinations of these functions. These combinations are simply the ones resulting from the transformation from the eigenbasis back to the standard basis. Denoting the elements of S by ϕ_j and those of S^- by ψ_j, we may now rewrite equation (B.49), restricted to a single Jordan block, in the form

$$Se^{D_{rs}t}e^{C_{rs}t}S^{-1} = e^{\lambda_r t} \sum_{p=1}^{d_{rs}} \sum_{q=1}^{d_{rs}} \frac{t^{q-p}}{(q-p)!} \phi_{ip} \psi_{qj} \sum_{k=1}^{q} \delta_{pk} \tag{B.57}$$

but the sum over δ simply selects terms for which $q \geq p$, so

that, by inverting the order of the summations and reindexing, we obtain

$$Se^{D_{rs}t}e^{C_{rs}t}S^{-1} = e^{\lambda_r t}\sum_{q=0}^{d_{rs}-1}\frac{t^q}{q!}\sum_{p=1}^{d_{rs}-q}\phi_{ip}\psi_{p+q\,j} \tag{B.58}$$

Finally, placing the blocks in their proper positions in J and summing over r and s yields

$$e^{At} = \sum_{r=1}^{n} e^{\lambda_r t}\sum_{s=1}^{m_r}\sum_{q=0}^{d_{rs}-1}\frac{t^q}{q!}\sum_{p=1}^{d_{rs}-q}\phi_{i\,u_{rs}+p-1}\psi_{u_{rs}+p+q-1\,j} \tag{B.59}$$

Now, the term under the rightmost summation sign expresses the outer product of one of the generalized eigenvectors with one of the dual generalized eigenvectors. That is,

$$\phi_{i\,u_{rs}+p-1}\psi_{u_{rs}+p+q-1} = \Phi_{r\,s\,p}\,\Psi^T_{r\,s\,p+q} \tag{B.60}$$

This product is just the matrix of coefficients arising from the transformation of a single element of e^{Jt}. It expresses the contribution of that element to one column of T. The rightmost sum in equation (B.59) is thus a matrix

$$E_{r\,s\,q} = \sum_{p=1}^{d_{rs}-q}\Phi_{r\,s\,p}\,\Psi^T_{r\,s\,p+q} \tag{B.61}$$

which is the matrix of coefficients of all terms involving t which arise from the block J . The matrix E is thus a projection of the eigenfunction containing t onto the vector Φ . We may thus rewrite equation (B.59) in terms of these projections, obtaining the form

$$e^{At} = \sum_{r=1}^{n} e^{\lambda_r t} \sum_{s=1}^{m_r} \sum_{q=0}^{d_{rs}-1} \frac{t^q}{q!} E_{rsq} \tag{B.62}$$

Now, when the eigenvalues are complex, the elements of the eigenvectors will, in general, be complex as well. Thus, letting

$$E_{rsq} = F_{rsq} + iG_{rsq} \tag{B.63}$$

and recalling that

$$e^{(\alpha+i\omega)t} = e^{\alpha t}(\cos\omega t + i\sin\omega t) \tag{B.64}$$

we may cast equation (B.59) in its final form:

$$e^{A\frac{t}{\tau}} = \sum_{r=1}^{n} \sum_{s=1}^{m_r} \sum_{q=0}^{d_{rs}-1} \frac{t^q}{q!\tau^q} e^{\alpha_r \frac{t}{\tau}} (F_{rsq}\cos\omega_r \frac{t}{\tau} - G_{rsq}\sin\omega_r \frac{t}{\tau}) \tag{B.65}$$

A review of the foregoing argument will show that the inverse of T is obtained from this expression simply by replacing t by -t. The response of the nerve ring to a perturbation in the inputs from the sensory preprocessing system are thus given by making the substitution from equation (B.65) into equation (B.37). It is clear that this response must always be real valued, even when the eigenvalues and eigenvectors are complex.

With respect to the form of the eigenfunctions, there are six cases to consider. Each eigenvalue may be real or complex, and in either case, its real part may be positive, negative, or zero. When the eigenvalue is complex, the

eigenfunction is a sinusoid modulated by the eigenfunction corresponding to the real part of the eigenvalue. The first few eigenfunctions are shown in figures B.2 a, B.3 a, B.4 and B.6. The remainder of this section focuses on the general case in which the eigenvalues are complex.

Within each of its invariant subspaces, the response of the nerve ring will depend only on the component of the input which lies in that subspace. This response consists of a superposition of sinusoids, each oscillating at angular frequency $\frac{\omega_r}{\tau}$. When $\alpha > 0$, one would expect the amplitude of the oscillations to increase without bound, but in this case the linearized analysis breaks down and the sigmoid transfer function of the neurons intervenes to stabilize the amplitude at some finite value. This leads to sustained oscillations of the type discussed in the next section. The effect of the sigmoidal nonlinearity on the eigenfunctions can be seen by comparing figures B.2 b, B.3 b, B.5 and B.7. to those cited above.

When $\alpha < 0$, on the other hand, the amplitude envelopes exhibit a series of peaks, evenly spaced at

$$t = -\frac{q\tau}{\alpha_r} \tag{B.66}$$

and by applying Stirling's formula,

$$q! \approx e^{-q} q^q \sqrt{2\pi q} \tag{B.67}$$

each peak in the sequence has an amplitude

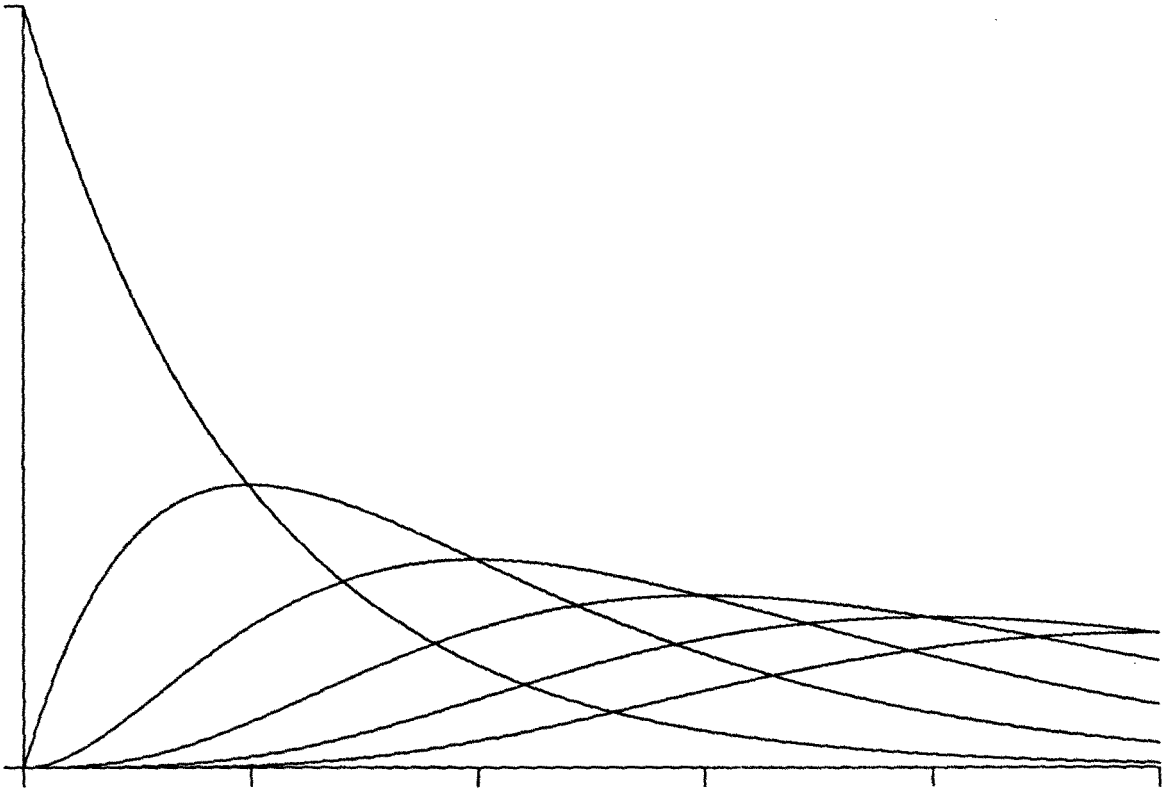

Figure B.2 a - The first six decreasing nonoscillatory eigenfunctions

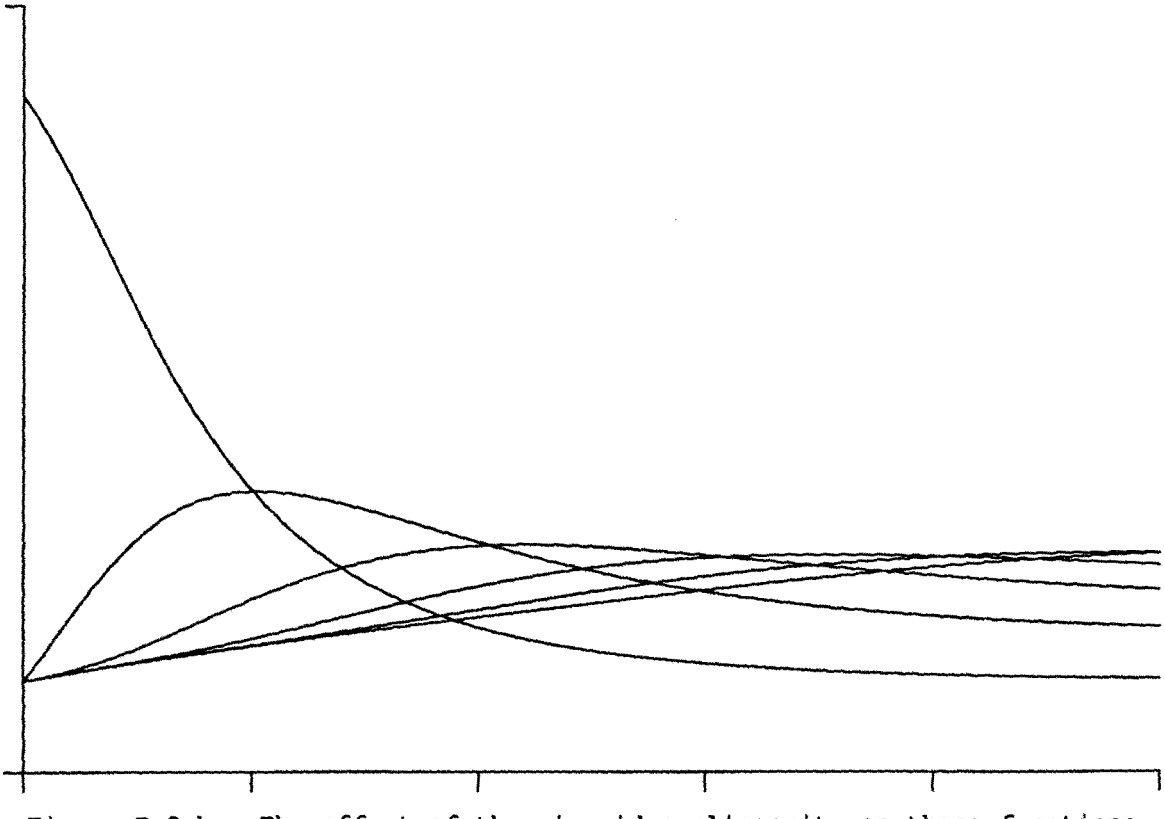

Figure B.2 b - The effect of the sigmoid nonlinearity on these functions

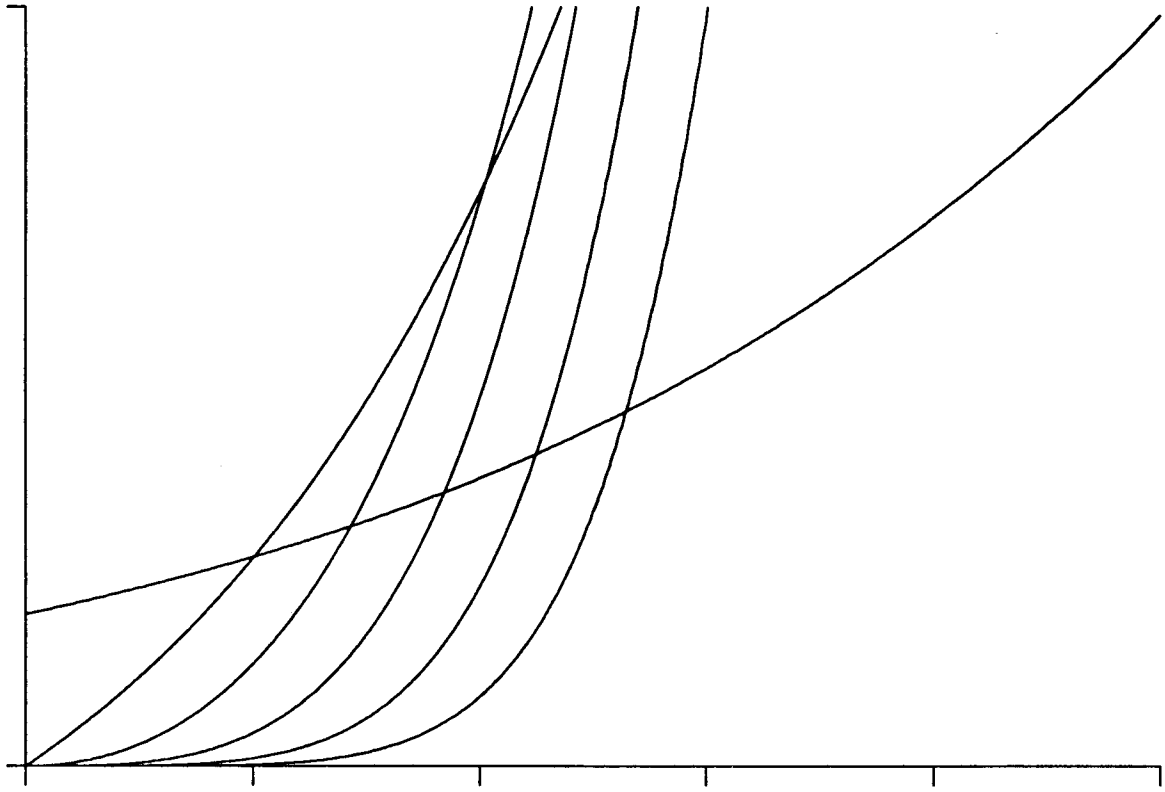

Figure B.3 a - The first six increasing nonoscillatory eigenfunctions

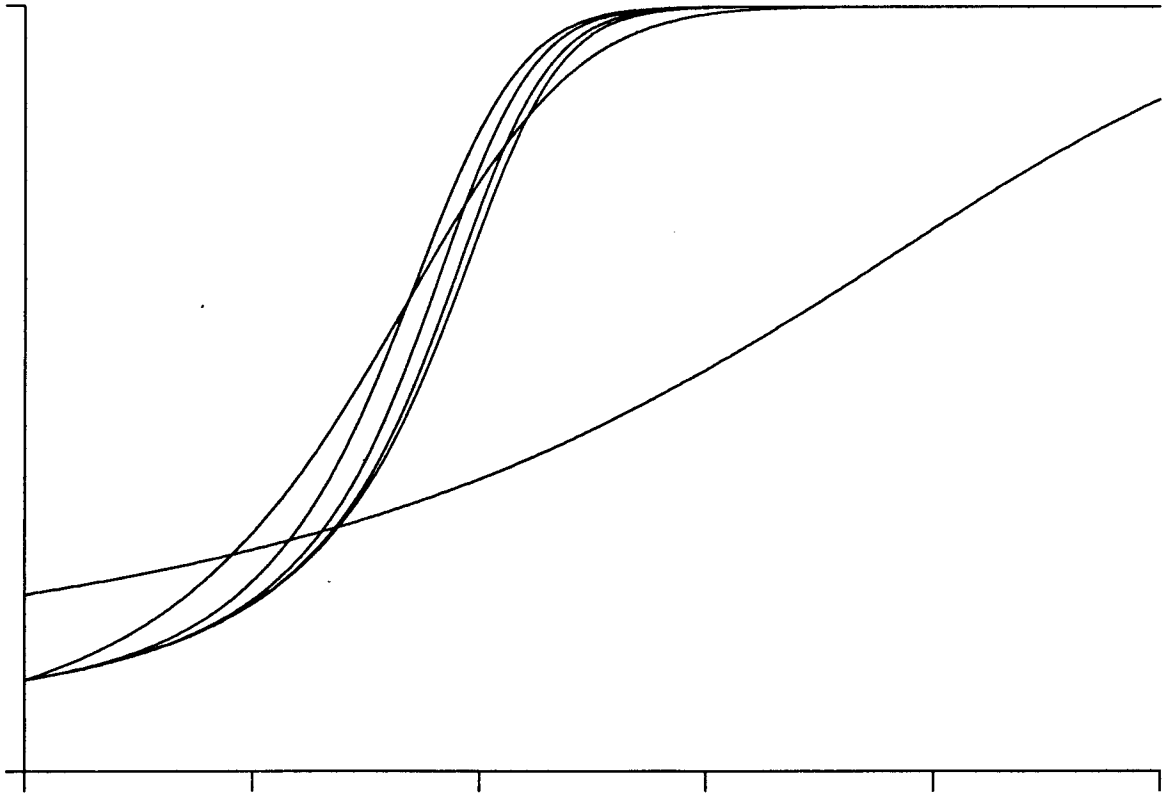

Figure B.3 b - The effect of the sigmoid nonlinearity on these functions

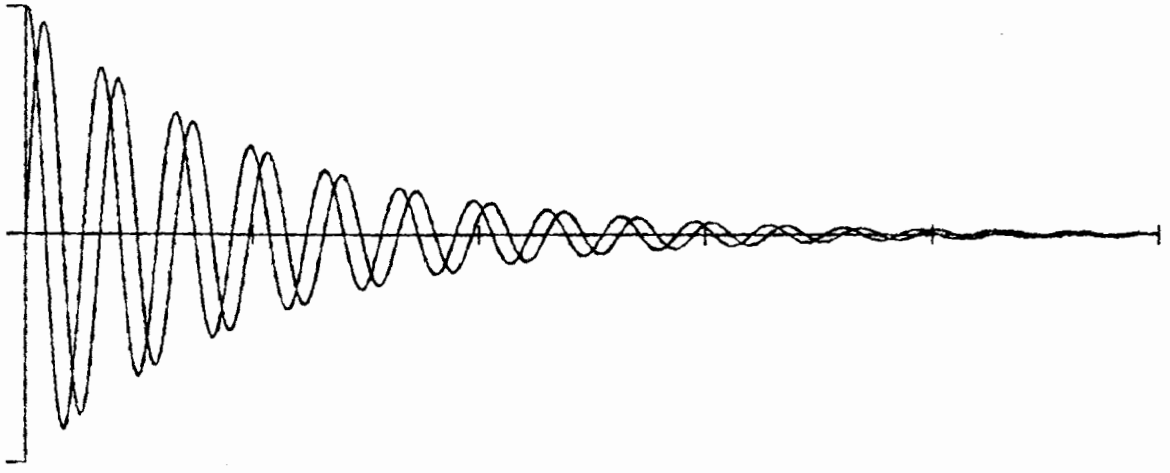

Figure B.4 a - Decreasing linear oscillatory eigenfunctions of index 0

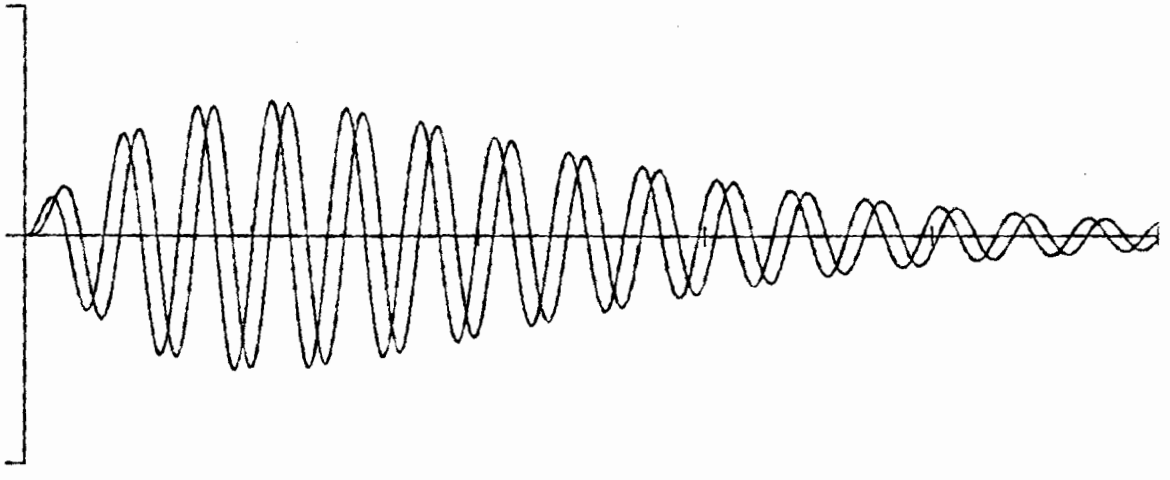

Figure B.4 b - Decreasing linear oscillatory eigenfunctions of index 1

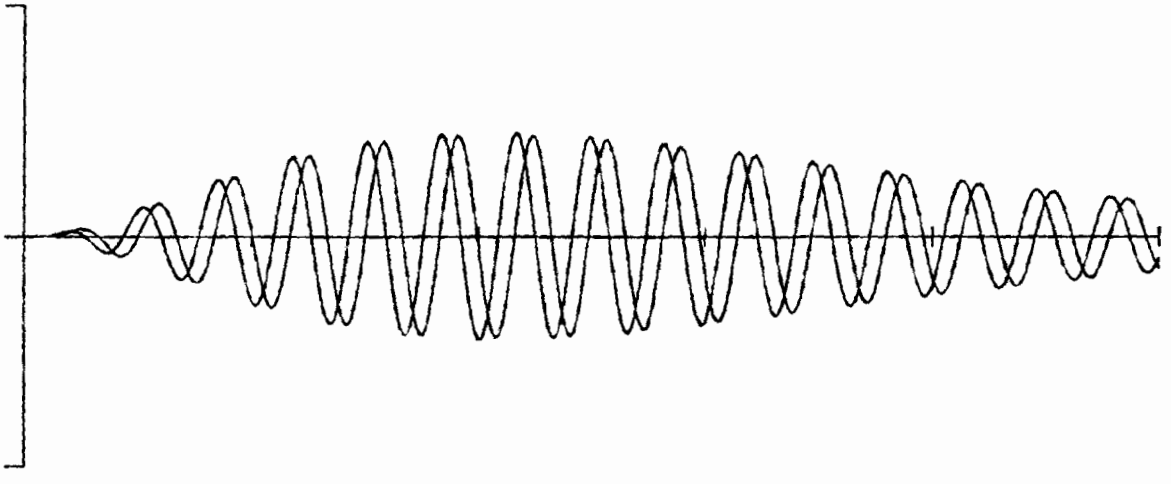

Figure B.4 c - Decreasing linear oscillatory eigenfunctions of index 2

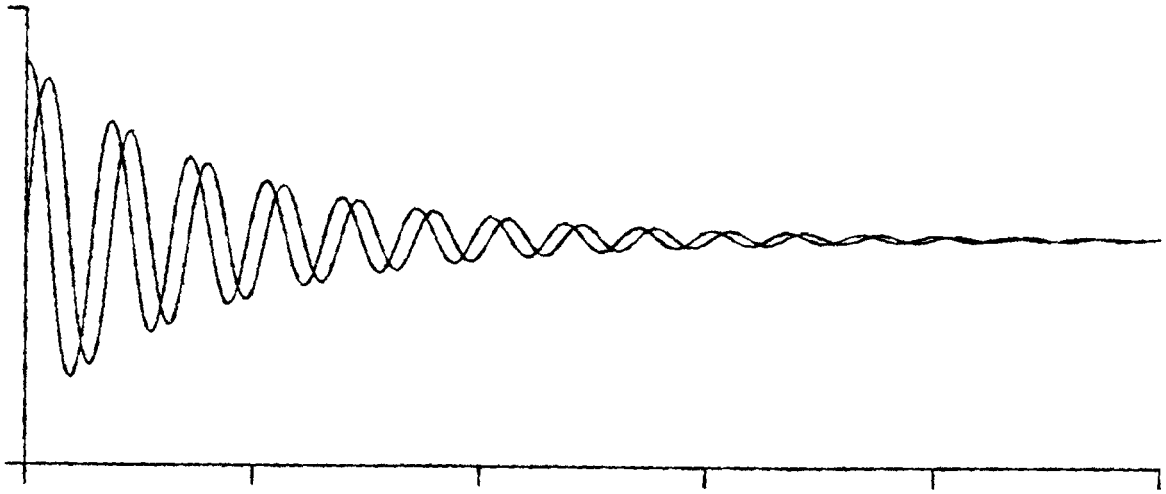

Figure B.5 a - Decreasing nonlinear oscillatory eigenfunctions, index 0

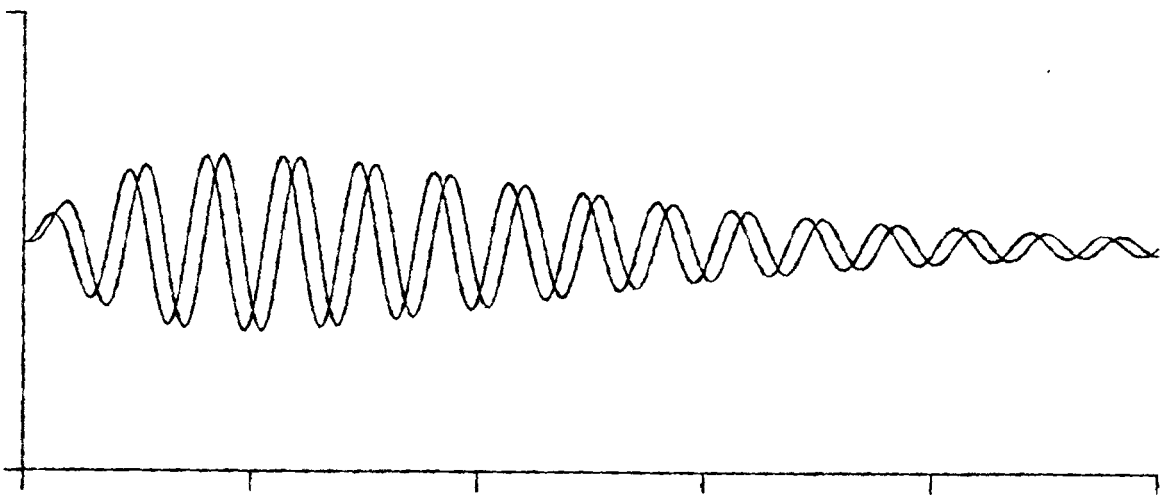

Figure B.5 b - Decreasing nonlinear oscillatory eigenfunctions, index 1

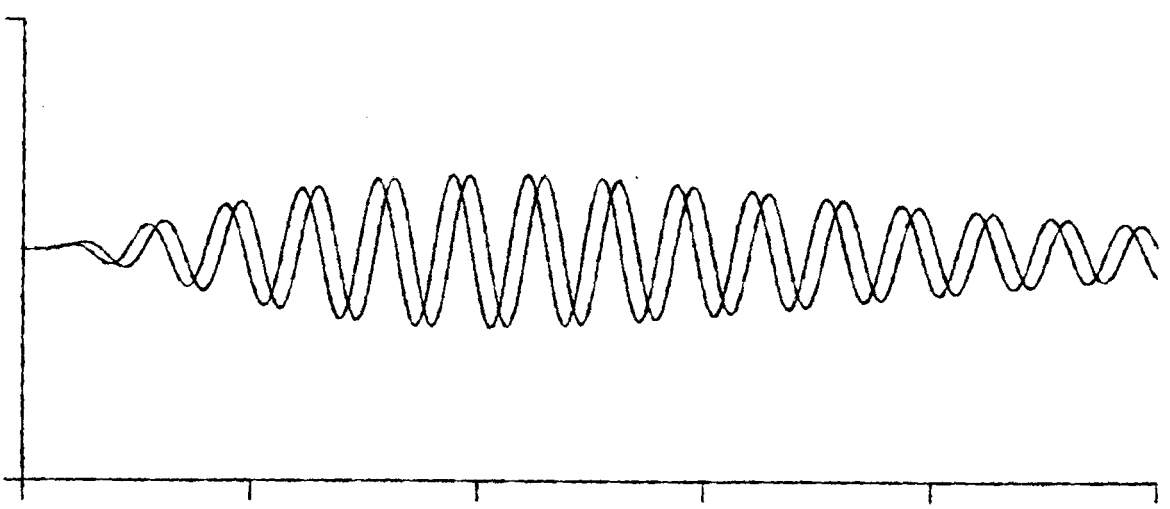

Figure B.5 c - Decreasing nonlinear oscillatory eigenfunctions, index 2

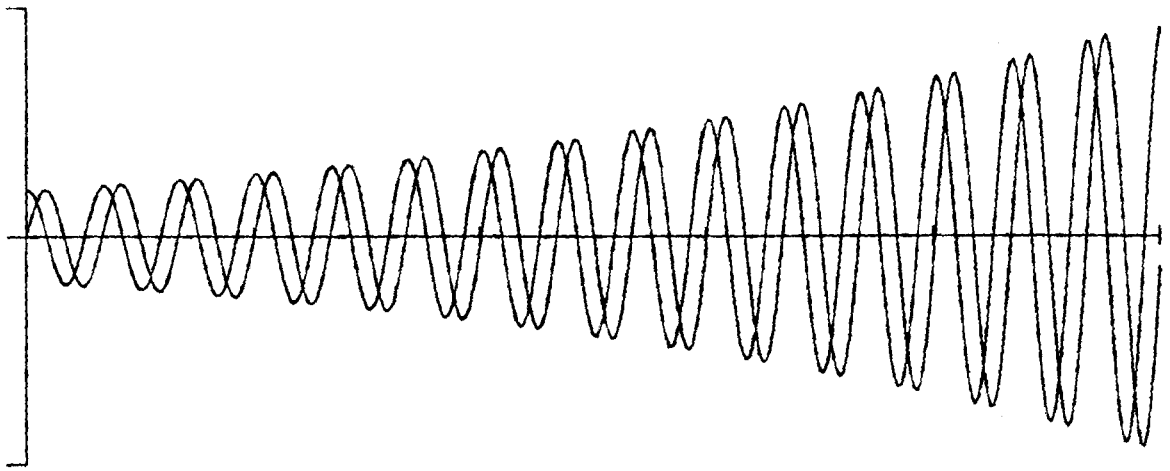

Figure B.6 a - Increasing linear oscillatory eigenfunctions of index 0

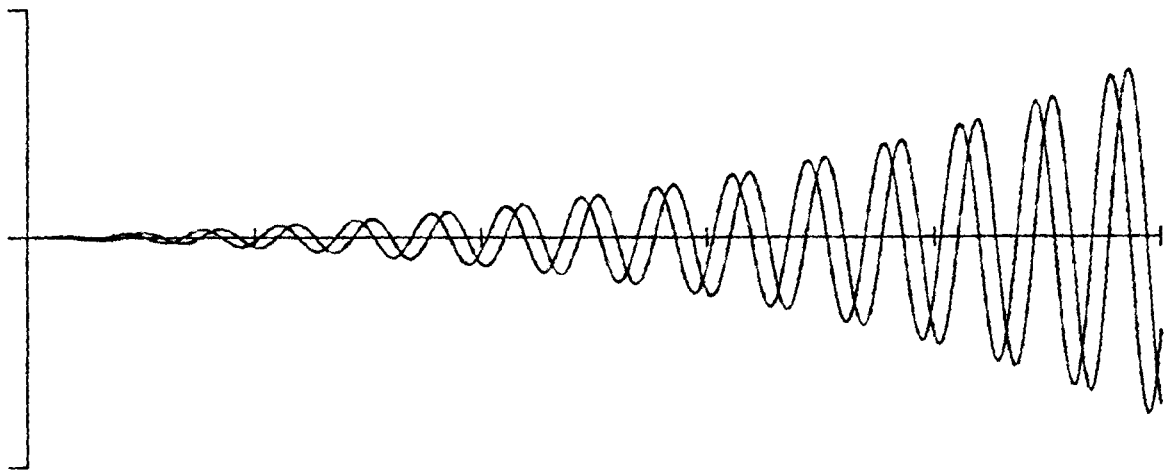

Figure B.6 b - Increasing linear oscillatory eigenfunctions of index 1

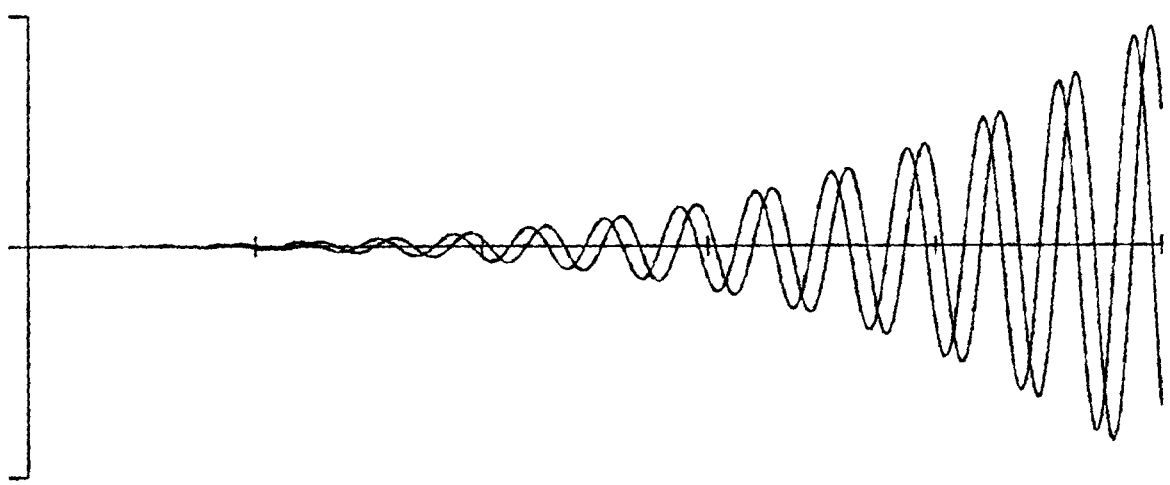

Figure B.6 c - Increasing linear oscillatory eigenfunctions of index 2

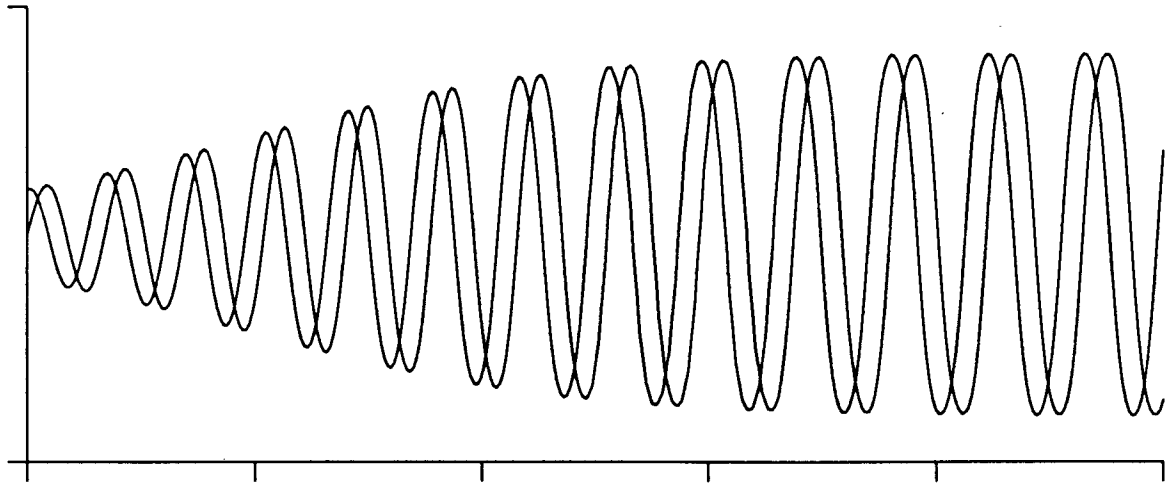

Figure B.7 a - Increasing nonlinear oscillatory eigenfunctions, index 0

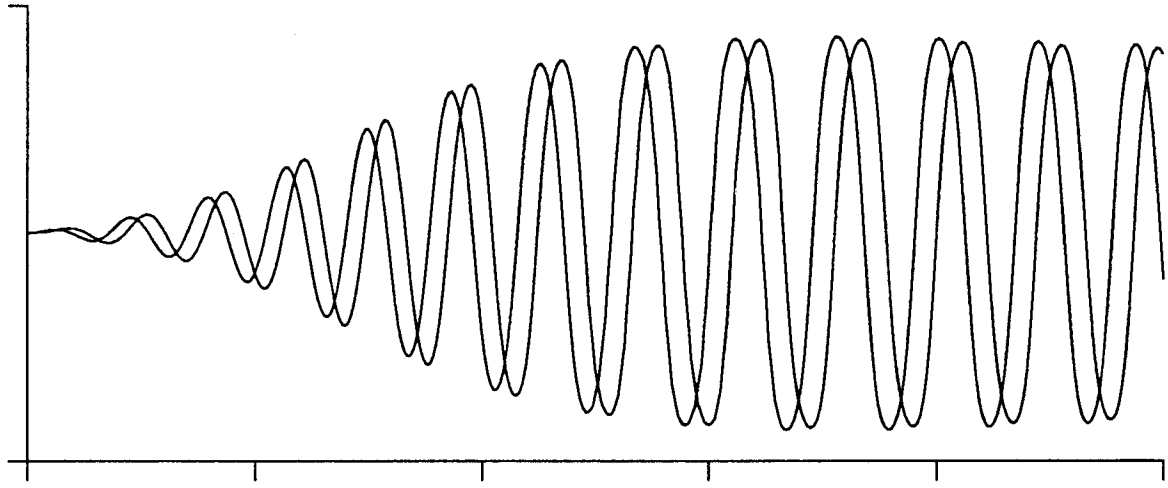

Figure B.7 b - Increasing nonlinear oscillatory eigenfunctions, index 1

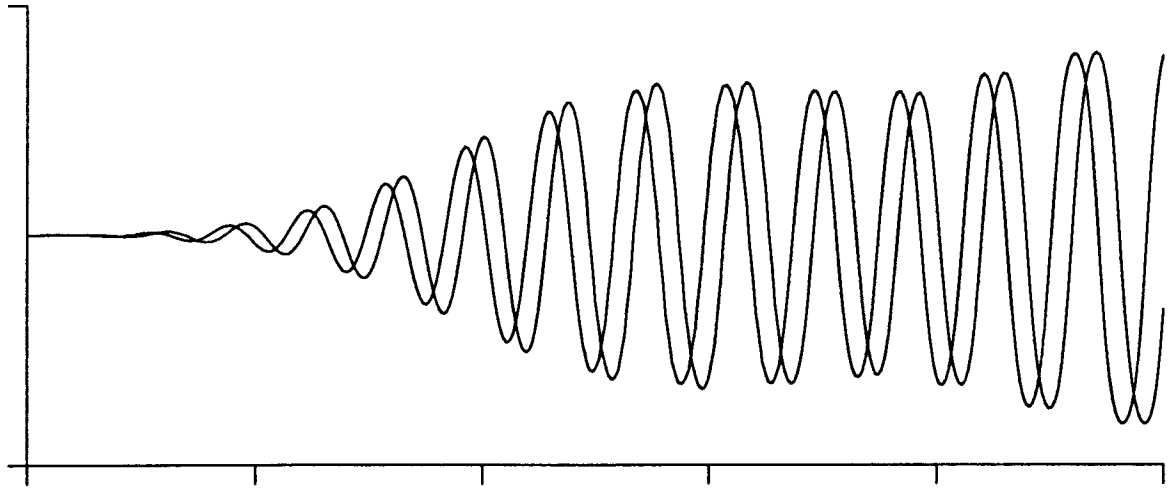

Figure B.7 c - Increasing nonlinear oscillatory eigenfunctions, index 2

$$\frac{t^q}{q!\tau^q} e^{\alpha_r \frac{t}{\tau}} \sim \frac{1}{\sqrt{2\pi q}(-\alpha_r)^q} \tag{B.68}$$

The number of peaks in the sequence is just d , the dimensionality of the subspace. The first peak occurs at t = 0 and lies parallel to Φ . Successive peaks lie parallel to Φ .

The nerve ring thus responds to a perturbation in the input from the sensory preprocessing system by issuing a series of motor commands. It does this in a manner which is both associative and context dependent. The context is determined by the interaction matrix, N, which picks out a set of critical neurons from the total complement of 69 neurons in the nerve ring. By choosing such a skeleton network out of the nerve ring, it determines the invariant subspaces of the ring at any given time, and thus controls the association of a particular sequence of motor commands with particular input from the sensory preprocessing system.

The conditions under which a context switch occurs are determined by the border matrix, B. The neurons picked out by this matrix are those for which a relatively small change in the input will cause an addition or deletion from the skeleton network picked out by the interaction matrix. An input lying parallel to a generalized eigenvector of B which is associated with a large eigenvalue will cause a variation in the structure of A. The matrix A can thus be thought of as a function of B. In this way, the nerve ring can associate different sequences of motor commands with the same input

signal, depending on the current background activity in the ring interneurons and the path through state space by which the system reached its current state.

The line of reasoning which lead through equation (B.14) to equation (B.28), the solution of which is obtained by substituting from equation (B.65) into equation (B.37), determines the dynamics of the membrane potentials under the assumption that the background activity is sufficiently stationary that the interaction matrix can be treated as constant. A similar line of reasoning would begin with equation (B.18) and determine the dynamics of the interaction matrix under the assumption that the border matrix can be treated as constant. While such an analysis is beyond the scope of this thesis, it would lead to a matrix differential equation, the solution of which would describe the interaction matrix as a function of time. This would lead to a characterization of the dynamics of the system in the regions of state space which have been excluded from the present analysis. These regions are the ones in which the dynamics of the membrane potentials undergo a bifurcation, giving rise to a change in the effective structure of the nerve ring effecting a context shift.

APPENDDIX C – CHARACTERIZATION OF TRAVELLING WAVES

A travelling sinusoidal wave is depicted in figure C.1, in which time increases to the left and distance along the body of the organism increases into the page. At each successive instant of time, the wave is displaced in space by a certain amount while retaining a constant waveform.

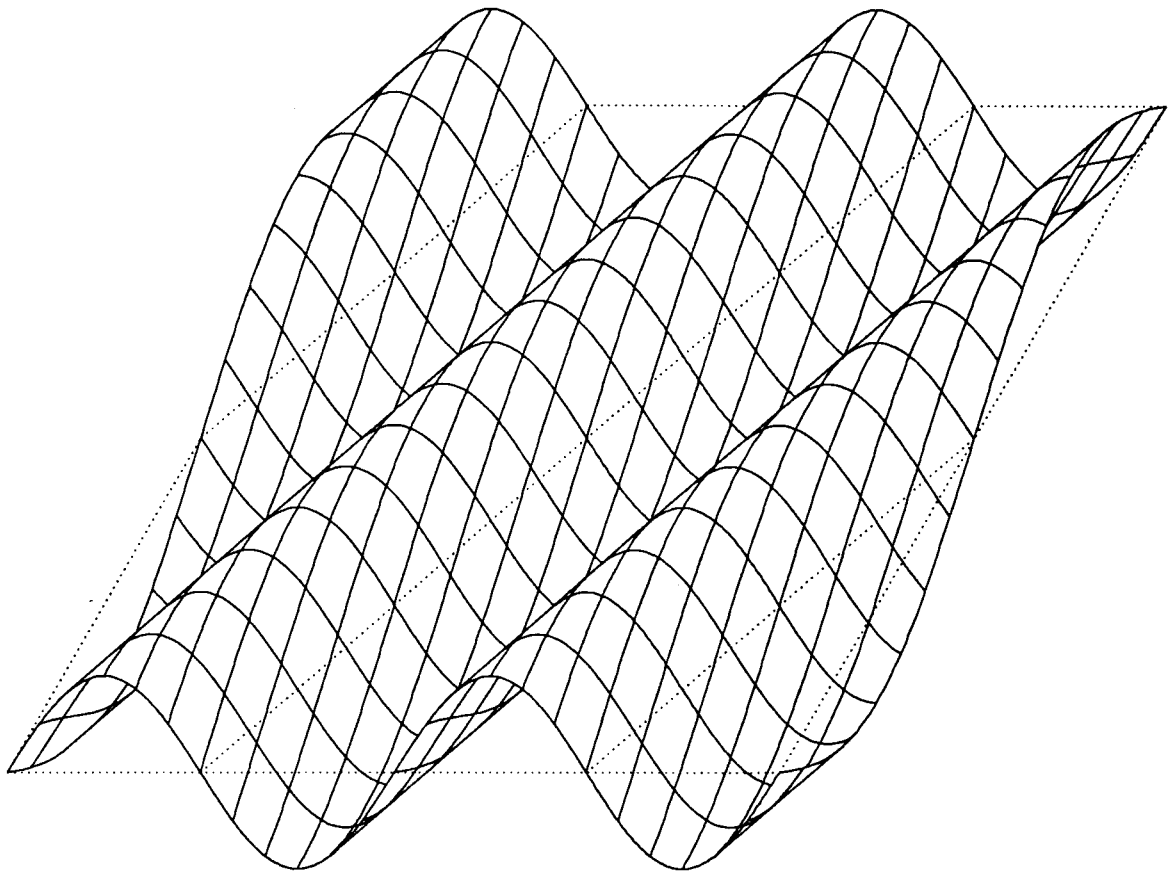

Figure C.1 - Wave propagation due to coupled oscillations

A relationship between the velocity of wave propagation, the frequency of the oscillations and the variation of phase shift with distance can be derived from figure C.2. In a time

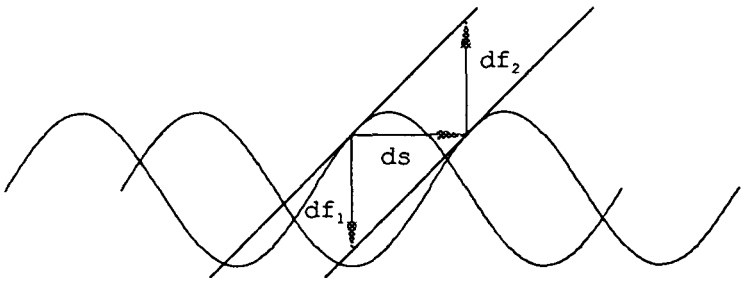

Figure C.2 - A travelling wave at successive instants

interval, dt, the wave has moved by distance, ds, so that

$$f(s+ds, t+dt) = f(s, t) \tag{C.1}$$

whereby

$$\frac{\partial}{\partial s} f(s+ds, t+dt) = \frac{\partial}{\partial s} f(s, t) \tag{C.2}$$

hence the two tangent lines in figure C.2 are parallel and

$$df_2 = -df_1 \tag{C.3}$$

But df_1 is just the change in f(s,t) over the interval dt, hence

$$df_1 = \frac{\partial}{\partial t} f(s, t)\, dt \tag{C.4}$$

and df_2 is the change in f(s,t) over displacement, ds, so that

$$df_2 = \frac{\partial}{\partial s} f(s, t)\, ds \tag{C.5}$$

Combining these last three expressions gives

$$\frac{\partial}{\partial t} f(s,t) = -\frac{ds}{dt}\frac{\partial}{\partial s} f(s,t) \tag{C.6}$$

but $\frac{ds}{dt}$ is nothing other than v, the velocity of propagation of the wave. If the wave is sinusoidal, having the form

$$f(s,t) = a\sin\theta(s,t) \tag{C.7}$$

then equation (C.6) has the form

$$\frac{\partial\theta}{\partial t} = -v\frac{\partial\theta}{\partial s} \tag{C.8}$$

and if the angular frequency,

$$\frac{\partial\theta}{\partial t} = \omega \tag{C.9}$$

is constant, then equation (C.8) yields

$$\frac{\partial\theta}{\partial s} = -\frac{\omega}{v} \tag{C.10}$$

which is constant as well. Letting

$$\frac{\partial\theta}{\partial s} = \alpha \tag{C.11}$$

be the rate at which the phase angle changes with distance along the body, the velocity of propagation is given by equation (C.10) as

$$v = -\frac{\omega}{\alpha} \tag{C.12}$$

and, because of equations (C.10) and (C.11), the phase angle must have the form

$$\theta(s,t) = \omega t + \alpha s + \theta_0 \tag{C.13}$$

hence

$$f(s,t) = a\sin(\omega t + \alpha s + \theta_0) \tag{C.14}$$

Now, the arguments leading up to equation (C.6) were independent of the waveform. All that was required was that the waveform remains invariant as the wave travels, as expressed by equation (C.1). Because any periodic function can be written as a sum of sinusoidally oscillating terms, and each term of the expansion must then satisfy equation (C.6), we must have

$$f(s,t) = \sum_{n=0}^{\infty} a_n \sin(n(\omega t + \alpha s) + \theta_n) \tag{C.15}$$

in order for equation (C.6) to be satisfied.

It is important to note that the variable, s, represents distance along the body of the worm, and the function f(s,t) represents the curvature of the body at time t and distance s from the tip of the head. Determination of the posture of the animal, in coordinates fixed to the ground, requires the evaluation of a non-elementary elliptic integral and, not being central to the issue, will not be pursued here.

The organism moves most efficiently in films of water having a depth on the order of half the diameter of the body.

A meniscus forms at the interface between the surface of the water film and the body of the worm. Distortion of the meniscus, due to flexion of the body, gives rise to a restoring force which, with increasing flexion, eventually becomes sufficient to prevent further side slip. Under these circumstances, the trajectory of the travelling wave will be fixed with respect to the ground and the worm will move along it at the velocity, v, derived above.

Appendix D - Solving for the Means and Variations

The equation of motion of the subsystem that mediates forward motion, equation (38), will be solved below in several steps. First, the the subsystem will be expressed in terms of the sums and differences of the homologous dorsoventrally opposed signals. This will show that the subsystem is itself composed of two decoupled subsystems. Next, each of these reduced subsystems is solved explicitly in its own generalized eigenbasis. Finally, the solutions of the reduced subsystems are recombined in the main text and transformed back into the basis formed by the linearized state variables of the subsystem.

To begin with, recall that equation (38) is a linear vector differential equation which reads

$$t\frac{d}{dt}\eta + A\eta = \xi \tag{D.1}$$

and let $\eta = Vx$. Then

$$t\frac{d}{dt}Vx + AVx = \xi \tag{D.2}$$

In this equation, the matrix A appears postmultiplied by V, and premultiplying the equation by V^{-} yields

$$t\frac{d}{dt}x + V^{-1}AVx = V^{-1}\xi \tag{D.3}$$

Explicitly, one has

$$V^{-1} = \begin{bmatrix} 1 & \cdot & \cdot & \cdot & \cdot & \cdot & 1 \\ \cdot & 1 & \cdot & \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & 1 & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & -1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot & -1 & \cdot & \cdot \\ \cdot & 1 & \cdot & \cdot & \cdot & -1 & \cdot \\ 1 & \cdot & \cdot & \cdot & \cdot & \cdot & -1 \end{bmatrix}, \quad V = \begin{bmatrix} \frac{1}{2} & \cdot & \cdot & \cdot & \cdot & \cdot & \frac{1}{2} \\ \cdot & \frac{1}{2} & \cdot & \cdot & \cdot & \frac{1}{2} & \cdot \\ \cdot & \cdot & \frac{1}{2} & \cdot & \frac{1}{2} & \cdot & \cdot \\ \cdot & \cdot & \cdot & -1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \frac{1}{2} & \cdot & -\frac{1}{2} & \cdot & \cdot \\ \cdot & \frac{1}{2} & \cdot & \cdot & \cdot & -\frac{1}{2} & \cdot \\ \frac{1}{2} & \cdot & \cdot & \cdot & \cdot & \cdot & -\frac{1}{2} \end{bmatrix} \tag{D.4}$$

and equation (38) is thereby transformed into

$$\tau \frac{d}{dt} \begin{bmatrix} x_{DM} \\ x_{BM} \\ x_{EM} \\ x_L \\ x_{EV} \\ x_{BV} \\ x_{DV} \end{bmatrix} + \begin{bmatrix} 1 & -e & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ d & -c & 1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 & \frac{2}{f} & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 & -c & -d \\ \cdot & \cdot & \cdot & -b & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & -e & 1 \end{bmatrix} \begin{bmatrix} x_{DM} \\ x_{BM} \\ x_{EM} \\ x_L \\ x_{EV} \\ x_{BV} \\ x_{DV} \end{bmatrix} = \begin{bmatrix} \cdot \\ a \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} \tag{D.5}$$

where the $x_{\cdot}$ are the sums and the $x_{\cdot v}$ are the differences. This is just equation (39) again and, as stated in the main body of the thesis, it clearly shows that the time development of the means is completely decoupled from the time development of the variations.

Now, V^{-} and V transform to and from a basis composed of sums and differences of moving averages of stochastic signals. Thus the sums follow the means, because the expected value of the random fluctuations vanishes and the systematic variations about the means are decoupled from the means by virtue of equation (D.5). A similar argument leads to the conclusion

that the differences follow the variations. The detailed solutions of the two reduced subsystems are given below.

Time Evolution of the Means

Proceeding first to solve for the means, we have

$$\tau \frac{d}{dt}\begin{bmatrix} x_{DM} \\ x_{BM} \\ x_{EM} \end{bmatrix} + \begin{bmatrix} 1 & -e & \cdot \\ \cdot & 1 & \cdot \\ d & -c & 1 \end{bmatrix}\begin{bmatrix} x_{DM} \\ x_{BM} \\ x_{EM} \end{bmatrix} = \begin{bmatrix} \cdot \\ a \\ \cdot \end{bmatrix} \tag{D.6}$$

that is,

$$t\frac{d}{dt}\eta_M + A_M\eta_M = \xi \tag{D.7}$$

To reduce A to its Jordan canonical form, we compute the determinant, $|A - \lambda I|$. Setting $\alpha = 1 - \lambda$,

$$\begin{vmatrix} \alpha & -e & \cdot \\ \cdot & \alpha & \cdot \\ d & -c & \alpha \end{vmatrix} = \alpha\begin{vmatrix} \alpha & -e \\ \cdot & \alpha \end{vmatrix} = \alpha^3 \tag{D.8}$$

and setting this determinant to zero results in $(1 - \lambda) = 0$ so that $\lambda = 0$ is the triple eigenvalue of A. There is only one eigenvector associated with this eigenvalue, as can be seen by solving $(A - \lambda I)\phi = 0$:

$$\begin{bmatrix} \cdot & -e & \cdot \\ \cdot & \cdot & \cdot \\ d & -c & \cdot \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -ex_2 \\ \cdot \\ dx_1 - cx_2 \end{bmatrix} = \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \rightarrow \phi_1 = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \cdot \\ \cdot \\ u \end{bmatrix} \tag{D.9}$$

where u is arbitrary. We thus compute the generalized eigen-

vectors by solving $(A - \lambda I)\phi = \phi_-$ as follows:

$$\begin{bmatrix} -ex_2 \\ \cdot \\ dx_1 - cx_2 \end{bmatrix} - \begin{bmatrix} \cdot \\ \cdot \\ u \end{bmatrix} \rightarrow \phi_2 - \begin{bmatrix} \frac{u}{d} \\ \cdot \\ v \end{bmatrix}, \quad \begin{bmatrix} -ex_2 \\ \cdot \\ dx_1 - cx_2 \end{bmatrix} - \begin{bmatrix} \frac{u}{d} \\ \cdot \\ v \end{bmatrix} \rightarrow \phi_3 - \begin{bmatrix} \frac{v-cu}{d} \\ -\frac{u}{de} \\ w \end{bmatrix} \quad (D.10)$$

where v and w are arbitrary. Now, the matrix which transforms from the standard basis to the eigenbasis is just that matrix whose columns are the eigenvectors and the inverse of this matrix, which transforms back to the standard basis, is the matrix whose rows are the dual eigenvectors. Thus, choosing u = de and v = c to simplify the form of the eigenvectors, we note that the determinant of this matrix is independent of w. We are thus free to chose w = 0. This gives the transform matrix to within a constant. Diagonalizing this matrix gives its inverse to within a constant, and normalizing the result to have all unit elements along the main diagonal provides a condition on these two constants. One obtains, finally,

$$S_M S_M^{-1} - \begin{bmatrix} \cdot & 1 & \cdot \\ \cdot & \cdot & -\frac{1}{c} \\ d & \frac{c}{e} & \cdot \end{bmatrix} \begin{bmatrix} -\frac{c}{de} & \cdot & \frac{1}{d} \\ 1 & \cdot & \cdot \\ \cdot & -e & \cdot \end{bmatrix} - \begin{bmatrix} 1 & \cdot & \cdot \\ \cdot & 1 & \cdot \\ \cdot & \cdot & 1 \end{bmatrix} \quad (D.11)$$

The Jordan form of A is thus given by

$$J_M - S_M^{-1} A_M S_M - \begin{bmatrix} -\frac{c}{de} & \cdot & \frac{1}{d} \\ 1 & \cdot & \cdot \\ \cdot & -e & \cdot \end{bmatrix} \begin{bmatrix} 1 & -e & \cdot \\ \cdot & 1 & \cdot \\ d & -c & 1 \end{bmatrix} \begin{bmatrix} \cdot & 1 & \cdot \\ \cdot & \cdot & -\frac{1}{e} \\ d & \frac{c}{e} & \cdot \end{bmatrix} - \begin{bmatrix} 1 & 1 & \cdot \\ \cdot & 1 & 1 \\ \cdot & \cdot & 1 \end{bmatrix} \quad (D.12)$$

It is easily verified that

$$J_M^{-1} = \begin{bmatrix} 1 & -1 & 1 \\ \cdot & 1 & -1 \\ \cdot & \cdot & 1 \end{bmatrix} \tag{D.13}$$

and by equations (B.49), (B.51) and (B.56) of appendix B,

$$e^{-J_M \frac{t}{\tau}} = \begin{bmatrix} 1 & -\frac{t}{\tau} & \frac{t^2}{2\tau^2} \\ \cdot & 1 & -\frac{t}{\tau} \\ \cdot & \cdot & 1 \end{bmatrix} e^{-\frac{t}{\tau}} \tag{D.14}$$

Now, introducing $y = S^{-} \eta$ and $x = S^{-} \xi$, equation (D.7) has the form

$$\tau \frac{d}{dt} y_M + J_M y_M = x_M \tag{D.15}$$

when expressed in the eigenbasis. Introducing the integrating factor $e^{J_M \frac{t}{\tau}}$ gives

$$\tau e^{J_M \frac{t}{\tau}} \frac{d}{dt} y_M \quad \tau y_M \quad \tau \frac{d}{} \quad \tau \qquad \tau \tag{D.16}$$

which can be immediately integrated to yield

$$y_M = e^{-J_M \frac{t}{\tau}} y_0 + J_M^{-1} \left(I - e^{-J_M \frac{t}{\tau}} \right) x_M \tag{D.17}$$

And assuming the subsystem starts from rest, that is $\eta = 0$, this becomes

$$\eta_M = S_M J_M^{-1} \left(I - e^{-J_M \frac{t}{\tau}} \right) S_M^{-1} \xi_M \tag{D.18}$$

when expressed in the standard basis again. After carrying out the matrix operations indicated in this expression, one obtains the result given in equation (40).

Time Evolution of the Variations

Turning now to the variation, η_v, about the mean, η, equation (40) reduces to

$$\tau \frac{d}{dt} \begin{bmatrix} x_L \\ x_{EV} \\ x_{BV} \\ x_{DV} \end{bmatrix} + \begin{bmatrix} 1 & \frac{2}{f} & \cdot & \cdot \\ \cdot & 1 & -c & -d \\ -b & \cdot & 1 & \cdot \\ \cdot & \cdot & e & 1 \end{bmatrix} \begin{bmatrix} x_L \\ x_{EV} \\ x_{BV} \\ x_{DV} \end{bmatrix} = \begin{bmatrix} \cdot \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} \tag{D.19}$$

or just

$$t \frac{d}{dt} \eta_v + A_v \eta_v = 0 \tag{D.20}$$

where, it will be recalled, the slopes of the sigmoids are absorbed into the coefficients. These slopes, of course, are to be evaluated at the mean, η. Hence, by equation (41), excitation from the nerve ring is required in order for equation (44) to have a nontrivial solution. In the absence of such excitation, η decays to zero, the slopes of the sigmoids become nearly zero, and the system becomes effectively decoupled.

Proceeding as before, we compute the eigenvalues by setting $|A_v - \lambda I|$ to zero. Setting $\alpha = 1 - \lambda$,

$$|A_v - \lambda I| = \alpha^4 + \frac{2b}{f}(\alpha c + de) = 0 \tag{D.21}$$

but a quartic polynomial can always be factored into two quadratic polynomials, say

$$(\alpha^2 + r\alpha + \tfrac{1}{2}(r^2 - s^2))\,(\alpha^2 - r\alpha + \tfrac{1}{2}(r^2 - s^2)) = 0 \qquad \text{(D.22)}$$

where r and s are real. These two, in turn, can be factored into linear factors

$$(\alpha - \alpha_1)\,(\alpha - \alpha_2)\,(\alpha - \alpha_3)\,(\alpha - \alpha_4) = 0 \qquad \text{(D.23)}$$

where α and α are complex conjugates, as are α and α. In fact,

$$\begin{aligned}
\alpha_1 &= \tfrac{1}{2}r\left(-1 + i\sqrt{1 - 2\tfrac{s^2}{r^2}}\right) \\
\alpha_2 &= \tfrac{1}{2}r\left(-1 - i\sqrt{1 - 2\tfrac{s^2}{r^2}}\right) \\
\alpha_3 &= \tfrac{1}{2}r\left(+1 + i\sqrt{1 + 2\tfrac{s^2}{r^2}}\right) \\
\alpha_4 &= \tfrac{1}{2}r\left(+1 - i\sqrt{1 + 2\tfrac{s^2}{r^2}}\right)
\end{aligned} \qquad \text{(D.24)}$$

Equating coefficients between equations (D.21) and (D.22) leads to the pair of equations

$$s^2 = \frac{2bc}{rf}, \quad s^4 = r^4 + \frac{8bde}{f} \qquad \text{(D.25)}$$

whereby

$$r^6 + \frac{8bde}{f}r^2 - \frac{4b^2c^2}{f^2} = 0 \qquad \text{(D.26)}$$

This is a cubic equation in r. Since a cubic equation always

has at least one real root, we have

$$r = p\left(\frac{2bc}{f}\right)^{\frac{1}{3}}, \quad s^2 = \frac{1}{p}\left(\frac{2bc}{f}\right)^{\frac{2}{3}} \tag{D.27}$$

where

$$p = \sqrt{\left(\frac{1}{2}\right)^{\frac{1}{3}}\left((q+1)^{\frac{1}{3}} + (q-1)^{\frac{1}{3}}\right)}, \quad q = \sqrt{1 + \frac{256}{27}\left(\frac{f}{2bc}\right)\left(\frac{de}{c}\right)^3} \tag{D.28}$$

and it can be seen that p, r and s are real and positive. Substituting equation (D.27) into (D.24) gives

$$\begin{aligned}
\alpha_1 &= \frac{p}{2}\left(\frac{2bc}{f}\right)\left(-1 + i\sqrt{1 - \frac{2}{p^3}}\right) \\
\alpha_2 &= \frac{p}{2}\left(\frac{2bc}{f}\right)\left(-1 - i\sqrt{1 - \frac{2}{p^3}}\right) \\
\alpha_3 &= \frac{p}{2}\left(\frac{2bc}{f}\right)\left(+1 + i\sqrt{1 + \frac{2}{p^3}}\right) \\
\alpha_4 &= \frac{p}{2}\left(\frac{2bc}{f}\right)\left(+1 - i\sqrt{1 + \frac{2}{p^3}}\right)
\end{aligned} \tag{D.29}$$

hence α and α are always complex while α and α may be real. The latter will be complex if $p > 2$, which is approximately equivalent to

$$\left(\frac{de}{c}\right)^3 > \frac{2bc}{13f} \tag{D.30}$$

When this condition does not hold, the eigenvalues λ and λ are real and positive, in which case any small deviation from equilibrium will increase exponentially until each neuron in the system are either saturated or cut off. Fortunately, most

reasonable assumptions about relative synaptic strengths and other biological factors lead to the satisfaction of inequality (D.30). We may therefore define

$$\alpha = \frac{p}{2}\left(\frac{2bc}{f}\right)^{\frac{1}{3}}, \quad \omega_1 = \frac{\alpha}{\tau}\sqrt{1 - \frac{2}{p^3}}, \quad \omega_2 = \frac{\alpha}{\tau}\sqrt{1 + \frac{2}{p^3}} \tag{D.31}$$

and

$$\theta_1 = \tan^{-1}\sqrt{1 - \frac{2}{p^3}}, \quad \theta_2 = \tan^{-1}\sqrt{1 + \frac{2}{p^3}} \tag{D.32}$$

and

$$r_1 = r\left(1 - \frac{2}{p^3}\right), \quad r_2 = r\left(1 + \frac{2}{p^3}\right) \tag{D.33}$$

whereby

$$\begin{aligned} \alpha_1 &= -r_1\, cis(-\theta_1) \\ \alpha_2 &= -r_1\, cis(+\theta_1) \\ \alpha_3 &= +r_2\, cis(+\theta_2) \\ \alpha_4 &= +r_2\, cis(-\theta_2) \end{aligned} \tag{D.34}$$

and the eigenvalues are

$$\begin{aligned} \lambda_1 &= 1 + \alpha - i\omega_1\tau \\ \lambda_2 &= 1 + \alpha + i\omega_1\tau \\ \lambda_3 &= 1 - \alpha - i\omega_2\tau \\ \lambda_4 &= 1 - \alpha + i\omega_2\tau \end{aligned} \tag{D.35}$$

Since there are four distinct eigenvalues, there are four linearly independent eigenvectors, and each of the invariant subspaces is one dimensional. The eigenvectors may be found

by solving $(A_v - \lambda I)\phi = 0$. That is,

$$\begin{bmatrix} \alpha_i & \frac{2}{f} & \cdot & \cdot \\ \cdot & \alpha_i & -c & -d \\ -b & \cdot & \alpha_i & \cdot \\ \cdot & \cdot & -e & \alpha_i \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} - \begin{bmatrix} \alpha_i x_1 + \frac{2}{f} x_2 \\ \alpha_i x_2 - cx_3 - dx_4 \\ \alpha_i x_3 - bx_1 \\ \alpha_i x_4 - ex_3 \end{bmatrix} - \begin{bmatrix} \cdot \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} \rightarrow \phi_i = u_i \begin{bmatrix} 1 \\ -\frac{\alpha_i f}{2} \\ \frac{b}{\alpha_i} \\ \frac{be}{\alpha_i^2} \end{bmatrix} \qquad (D.36)$$

The dual eigenvectors are found by solving $\psi^T(A_v - \lambda I) = 0$. That is,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}^T \begin{bmatrix} \alpha_i & \frac{2}{f} & \cdot & \cdot \\ \cdot & \alpha_i & -c & -d \\ -b & \cdot & \alpha_i & \cdot \\ \cdot & \cdot & -e & \alpha_i \end{bmatrix} - \begin{bmatrix} \alpha_i x_1 - bx_3 \\ \alpha_i x_2 - \frac{2}{f} x_1 \\ \alpha_i x_3 - cx_2 - ex_4 \\ \alpha_i x_4 - dx_2 \end{bmatrix}^T - \begin{bmatrix} \cdot \\ \cdot \\ \cdot \\ \cdot \end{bmatrix}^T \rightarrow \psi_i^T = v_i \begin{bmatrix} 1 \\ -\frac{2}{\alpha_i f} \\ \frac{\alpha_i}{b} \\ -\frac{2d}{\alpha_i^2 f} \end{bmatrix}^T \qquad (D.37)$$

As before, the transform from the standard basis to the eigenbasis, S_v, is the matrix whose columns are the ϕ and its inverse is the matrix whose rows are the ψ^T. The elements of the product of these two matrices are just the inner products of these vectors. It is straight forward, though mechanically intricate, to show that

$$\psi_i^T \phi_j = \left(3 - \frac{2bde}{\alpha_i^4 f}\right) u_j v_i \delta_{ij} \qquad (D.38)$$

Hence we obtain a condition on u and v :

$$\left(3 - \frac{2bde}{\alpha_i^4 f}\right) u_i v_i = 1 \qquad (D.39)$$

but from equation (D.21),

$$\alpha_i^4 = -\frac{2b}{f}(\alpha_i c + de) \tag{D.40}$$

and thus, after some manipulation, equation (D.39) can be made to read

$$u_i v_i = \frac{-\alpha_i^4\left(4\left(\frac{2bde}{f}\right) + 3\left(\frac{2bc}{f}\right)\alpha_i^*\right)}{16\left(\frac{2bde}{f}\right)^2 - 9\left(\frac{2bc}{f}\right)^2 + 12\left(\frac{2bde}{f}\right)\left(\frac{2bc}{f}\right)(\alpha_i + \alpha_i^*)} \tag{D.41}$$

in which the denominator is real.

We may therefore define $u = \alpha$ and $v = \alpha\beta$, where β is what remains of the right hand side of equation (D.41) when α has been factored out. In the following argument, it will be useful to have the β in polar form. Now, it is easy to verify that β and β are complex conjugates, as are β and β. We may therefore let

$$\begin{aligned}
\beta_1 &= r_3\, cis(+\theta_3)\\
\beta_2 &= r_3\, cis(-\theta_3)\\
\beta_3 &= r_4\, cis(+\theta_4)\\
\beta_4 &= r_4\, cis(-\theta_4)
\end{aligned} \tag{D.42}$$

and write

$$S_V = \begin{bmatrix} \alpha_1^2 & \alpha_2^2 & \alpha_3^2 & \alpha_4^2 \\ -\frac{\alpha_1^3 f}{2} & -\frac{\alpha_2^3 f}{2} & -\frac{\alpha_3^3 f}{2} & -\frac{\alpha_4^3 f}{2} \\ \alpha_1 b & \alpha_2 b & \alpha_3 b & \alpha_4 b \\ be & be & be & be \end{bmatrix}, \quad S_V^{-1} = \begin{bmatrix} \alpha_1^2\beta_1 & -\frac{2\alpha_1\beta_1}{f} & \frac{\alpha_1^3\beta_1}{b} & -\frac{2d\beta_1}{f} \\ \alpha_2^2\beta_2 & -\frac{2\alpha_2\beta_2}{f} & \frac{\alpha_2^3\beta_2}{b} & -\frac{2d\beta_2}{f} \\ \alpha_3^2\beta_3 & -\frac{2\alpha_3\beta_3}{f} & \frac{\alpha_3^3\beta_3}{b} & -\frac{2d\beta_3}{f} \\ \alpha_4^2\beta_4 & -\frac{2\alpha_4\beta_4}{f} & \frac{\alpha_4^3\beta_4}{b} & -\frac{2d\beta_4}{f} \end{bmatrix} \tag{D.43}$$

These expressions can be used to verify that the Jordan form of A_v is simply

$$J_V = \begin{bmatrix} \lambda_1 & \cdot & \cdot & \cdot \\ \cdot & \lambda_2 & \cdot & \cdot \\ \cdot & \cdot & \lambda_3 & \cdot \\ \cdot & \cdot & \cdot & \lambda_4 \end{bmatrix}, \quad J_V^{-1} = \begin{bmatrix} \frac{1}{\lambda_1} & \cdot & \cdot & \cdot \\ \cdot & \frac{1}{\lambda_2} & \cdot & \cdot \\ \cdot & \cdot & \frac{1}{\lambda_3} & \cdot \\ \cdot & \cdot & \cdot & \frac{1}{\lambda_4} \end{bmatrix} \tag{D.44}$$

and

$$e^{-J_V\frac{t}{\tau}} = \begin{bmatrix} e^{-\lambda_1\frac{t}{\tau}} & \cdot & \cdot & \cdot \\ \cdot & e^{-\lambda_2\frac{t}{\tau}} & \cdot & \cdot \\ \cdot & \cdot & e^{-\lambda_3\frac{t}{\tau}} & \cdot \\ \cdot & \cdot & \cdot & e^{-\lambda_4\frac{t}{\tau}} \end{bmatrix} \tag{D.45}$$

We may now proceed to give the solution to equation (D.20). It is obtained as before, by transforming the equation to the eigenbasis, integrating using $e^{J_V\frac{t}{\tau}}$ as the integrating factor and transforming back to the standard basis. Since the equation is homogeneous, the result is

$$\eta_V = S_V e^{-J_V\frac{t}{\tau}} S_V^{-1} \eta_0 \tag{D.46}$$

From equations (D.43) and (D.45),

$$S_V e^{-J_V\frac{t}{\tau}} = \begin{bmatrix} \alpha_1^2 e^{-\lambda_1\frac{t}{\tau}} & \alpha_2^2 e^{-\lambda_1\frac{t}{\tau}} & \alpha_3^2 e^{-\lambda_1\frac{t}{\tau}} & \alpha_4^2 e^{-\lambda_1\frac{t}{\tau}} \\ -\frac{\alpha_1^3 f}{2} e^{-\lambda_2\frac{t}{\tau}} & -\frac{\alpha_2^3 f}{2} e^{-\lambda_2\frac{t}{\tau}} & -\frac{\alpha_3^3 f}{2} e^{-\lambda_2\frac{t}{\tau}} & -\frac{\alpha_4^3 f}{2} e^{-\lambda_2\frac{t}{\tau}} \\ \alpha_1 b e^{-\lambda_3\frac{t}{\tau}} & \alpha_2 b e^{-\lambda_3\frac{t}{\tau}} & \alpha_3 b e^{-\lambda_3\frac{t}{\tau}} & \alpha_4 b e^{-\lambda_3\frac{t}{\tau}} \\ b e e^{-\lambda_4\frac{t}{\tau}} & b e e^{-\lambda_4\frac{t}{\tau}} & b e e^{-\lambda_4\frac{t}{\tau}} & b e e^{-\lambda_4\frac{t}{\tau}} \end{bmatrix} \tag{D.47}$$

and the product $S_V^{-}\eta$ can be written as the sum of four vectors, as follows.

$$S_V^{-1}\eta_0 = \begin{bmatrix} \alpha_1^2\beta_1 \\ \alpha_2^2\beta_2 \\ \alpha_3^2\beta_3 \\ \alpha_4^2\beta_4 \end{bmatrix} x_{LO} - \frac{2}{f}\begin{bmatrix} \alpha_1\beta_1 \\ \alpha_2\beta_2 \\ \alpha_3\beta_3 \\ \alpha_4\beta_4 \end{bmatrix} x_{EVO} + \frac{1}{b}\begin{bmatrix} \alpha_1^2\beta_1 \\ \alpha_2^2\beta_2 \\ \alpha_2^2\beta_2 \\ \alpha_4^2\beta_4 \end{bmatrix} x_{BVO} - \frac{2d}{f}\begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} x_{DVO} \qquad (D.48)$$

It now remains to take the product of the last two equations but, reasoning that the expected initial values of the variations of the neuron signals and muscular excitation are zero, we will consider only the effect of the initial posture of the organism. Thus

$$\eta_V = \begin{bmatrix} \alpha_1^4\beta_1 e^{-\lambda_1\frac{t}{\tau}} + \alpha_2^4\beta_2 e^{-\lambda_2\frac{t}{\tau}} + \alpha_3^4\beta_3 e^{-\lambda_3\frac{t}{\tau}} + \alpha_4^4\beta_4 e^{-\lambda_4\frac{t}{\tau}} \\ -\frac{f}{2}(\alpha_1^5\beta_1 e^{-\lambda_1\frac{t}{\tau}} + \alpha_2^5\beta_2 e^{-\lambda_2\frac{t}{\tau}} + \alpha_3^5\beta_3 e^{-\lambda_3\frac{t}{\tau}} + \alpha_4^5\beta_4 e^{-\lambda_4\frac{t}{\tau}}) \\ b(\alpha_1^3\beta_1 e^{-\lambda_3\frac{t}{\tau}} + \alpha_2^3\beta_2 e^{-\lambda_3\frac{t}{\tau}} + \alpha_3^3\beta_3 e^{-\lambda_3\frac{t}{\tau}} + \alpha_4^3\beta_4 e^{-\lambda_3\frac{t}{\tau}}) \\ be(\alpha_1^2\beta_1 e^{-\lambda_4\frac{t}{\tau}} + \alpha_2^2\beta_2 e^{-\lambda_4\frac{t}{\tau}} + \alpha_3^2\beta_3 e^{-\lambda_4\frac{t}{\tau}} + \alpha_4^2\beta_4 e^{-\lambda_4\frac{t}{\tau}}) \end{bmatrix} X_{LO} \qquad (D.49)$$

and so we proceed to investigate the sum of terms of the form $\alpha_i^k\beta_i e^{-\lambda_i\frac{t}{\tau}}$. We know that

$$e^{r+i\theta} = e^r cis\theta \qquad (D.50)$$

so that, from equation (D.35), we have

$$\begin{aligned} e^{-\lambda_1\frac{t}{\tau}} &= e^{-(\alpha+1)\frac{t}{\tau}} cis(+\omega_1 t) \\ e^{-\lambda_2\frac{t}{\tau}} &= e^{-(\alpha+1)\frac{t}{\tau}} cis(-\omega_1 t) \\ e^{-\lambda_3\frac{t}{\tau}} &= e^{(\alpha-1)\frac{t}{\tau}} cis(+\omega_2 t) \\ e^{-\lambda_4\frac{t}{\tau}} &= e^{(\alpha-1)\frac{t}{\tau}} cis(-\omega_2 t) \end{aligned} \qquad (D.51)$$

and using these expressions together with equation (D.34) and (D.41) gives

$$\begin{aligned} \alpha_1^k \beta_1 e^{-\lambda_1 \frac{t}{\tau}} &= (-r_1)^k r_3 e^{-(\alpha+1)\frac{t}{\tau}} cis(+\theta_3 - k\theta_1 + \omega_1 t) \\ \alpha_2^k \beta_2 e^{-\lambda_2 \frac{t}{\tau}} &= (-r_1)^k r_3 e^{-(\alpha+1)\frac{t}{\tau}} cis(-\theta_3 + k\theta_1 - \omega_1 t) \end{aligned} \tag{D.52}$$

and

$$\begin{aligned} \alpha_3^k \beta_3 e^{-\lambda_3 \frac{t}{\tau}} &= r_2^k r_4 e^{(\alpha-1)\frac{t}{\tau}} cis(+\theta_4 + k\theta_2 + \omega_2 t) \\ \alpha_4^k \beta_4 e^{-\lambda_4 \frac{t}{\tau}} &= r_2^k r_4 e^{(\alpha-1)\frac{t}{\tau}} cis(-\theta_4 - k\theta_2 - \omega_2 t) \end{aligned} \tag{D.53}$$

Now, the right hand sides of equations (D.52) have identical magnitudes and opposite phase angles, hence their sum is real. Likewise for equations (D.53). In fact,

$$cis\theta + cis(-\theta) = 2\cos\theta \tag{D.54}$$

whereby

$$\begin{aligned} \alpha_1^k \beta_1 e^{-\lambda_1 \frac{t}{\tau}} + \alpha_2^k \beta_2 e^{-\lambda_2 \frac{t}{\tau}} &= 2(-r_1)^k r_3 e^{-(\alpha+1)\frac{t}{\tau}} \cos(\omega_1 t + \theta_3 - k\theta_1) \\ \alpha_3^k \beta_3 e^{-\lambda_3 \frac{t}{\tau}} + \alpha_4^k \beta_4 e^{-\lambda_4 \frac{t}{\tau}} &= 2 r_2^k r_4 e^{(\alpha-1)\frac{t}{\tau}} \cos(\omega 2_t + \theta_4 + k\theta_2) \end{aligned} \tag{D.55}$$

and substituting these expressions into equation (D.48) allows one, finally, to write the solution to equation (D.20) assuming that only x has a nonzero initial value. The result is given by equation (45) in the main text. Now the first term on the right hand side of this equation is exponentially damped, so it will decay to zero over time. The second term will grow exponentially if $\alpha > 1$. From equation (D.31), this

amounts to

$$\frac{p}{2}\left(\frac{2bc}{f}\right)^{\frac{1}{3}} > 1 \qquad \text{(D.56)}$$

which is easily satisfied, given reasonable biological assumptions.

REFERENCES

Ackley, D H, Hinton, G E, and Sejnowski, T J, 1985:

"A Learning Algorithm for Boltzmann Machines",
Cognitive Science, Vol 9 Nbr 1 (1985)

Adomian, G, 1983:

Stochastic Systems, Academic Press (1983)

Albertson, D G, and Thomson, J N, 1976:

"The Pharynx of *Caenorhabditis elegans*",
Phil. Trans. R. Soc. London (Biol), 275: 299 - 325 (1976)

an der Heiden, U, 1980:

Analysis of Neural Networks, Springer-Verlag (1980)

Caianiello, E R, de Luca, A, and Ricciardi, L M, 1967:

"Neural Networks: Reverberations, Constants of Motion, General Behaviour",
Proceedings of the Ravello School on Neural Networks,
Ed. E R Caianiello, Springer-Verlag (1967)

Chalfie, M, Sulston, J E, White, J G, Southgate, E,
Thomson, J N, and Brenner, S, 1985:

"The Neural Circuit for Touch Sensitivity in *Caenorhabditis elegans*",
Journal of Neuroscience, Vol 5 Nbr 4 (1985)

Chalfie, M, and White, J G, 1988:

"The Nervous System", The Nematode *Caenorhabditis elegans*,
Ed. W B Wood, Cold Spring Harbor Laboratory (1988)

Cox, R T, 1978:

"Of Inference and Inquiry, An Essay in Inductive Logic",
The Maximum Entropy Formalism,
Ed. R D Levine and M Tribus, MIT Press (1978)

Crofton, H D, 1966:

Nematodes, Hutchinson & Co, London (1966)

Croll, N A, 1970:

The Behaviour of Nematodes, St. Martin's, New York (1970)

Fodor, J A, and Pylyshyn, Z W, 1988:

"Connectionism & cognitive architecture: A critical analysis",
Cognition, Vol 28 (1988), pp 3 - 71

Geman, S, and Geman, D, 1984:

"Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images",
IEEE Trans. Pattern Analysis and Machine Intelligence,
Vol 7 Nbr 1 (1984)

Hall, D H, 1977:

The Posterior Nervous System of the Nematode *Caenorhabditis elegans*,
Ph.D. Thesis, California Institute of Technology (1977)

Harth, E, 1982:

"On the Spontaneous Emergence of Neuronal Schemata",
Competition and Cooperation in Neural Nets,
Ed. S Amari and M A Arbib, Springer-Verlag (1982)

Hinton, G E, and Sejnowski, T J, 1983:

"Optimal Perceptual Inference", Proc. IEEE Comp. Soc. Conf. Computer Vision and Pattern Recognition (1983)

Honig, M L, and Messerschmitt, D G, 1984:

Adaptive Filters: Structures, Algorithms, Applications
Kluwer Academic Publishers, Boston (1984)

Hopfield, J J, 1982:

"Neural Networks and Physical Systems with Emergent Collective Computational Abilities",
Proc. Natl. Acad. Sci. USA, Vol 79 (April 1982)

Isham, V, 1981:

"An Introduction to Spacial Point Processes and Markov Random Fields",
International Statistical Review, Vol 49 (1981)

Jaynes, E T, 1957:

"Information Theory and Statistical Mechanics"
Physical Review, Vol 106: 620 - 630 (1957)

Jaynes, E T, 1978:

"Where Do We Stand on Maximum Entropy?",
The Maximum Entropy Formalism,
Ed. R D Levine and M Tribus, MIT Press (1978)

Johnson, C D, and Stretton, A O W, 1980:

"Neural Control of Locomotion in *Ascaris*: Anatomy, Electrophysiology, and Biochemistry",
Nematodes as Biological Models, Volume 1,
Ed. B M Zuckerman, Academic Press, New York (1980)

Kristan, W B, Jr, 1980:

"Generation of Rhythmic Motor Patterns",
Information Processing in the Nervous System,
Ed. H M Pinsker and W D Willis Jr, Raven Press (1980)

Kohonen, T, 1984:

Self-Organization and Associative Memory,
Springer-Verlag (1984)

Kullback, C, 1968:

Information Theory and Statistics, Dover Books (1968)

Lee, D L, 1965:

The Physiology of Nematodes, W. H. Freeman and Company, San Francisco (1984)

Mill, P J, 1982:

Comparative Neurobiology, Edward Arnold, London (1982)

Nicolis, G, and Prigogine, I, 1977:

Self-Organization in Nonequilibrium Systems, John Wiley and Sons (1977)

Palm, G, 1982:

Neural Assemblies, Springer-Verlag (1982)

Sejnowski, T J, 1976:

"On the Stochastic Dynamics of Neuronal Interaction", Biol. Cybernetics, 22: 203 - 211 (1976)

Wannier, G H, 1966:

Statistical Physics, John Wiley and Sons (1966)

Ward, S, 1976:

"The use of mutants to analyze the sensory nervous system of *Caenorhabditis elegans*", The Organization of Nematodes, Ed. N A Croll, Academic Press (1976)

Ward, S, Thomson, J N, White, J G, and Brenner, S, 1975:

"Electron Microscopical Reconstruction of the Anterior Sensory Anatomy of the Nematode *Caenorhabditis elegans*", Journal of Comparative Neurology, 160: 313 - 338 (1975)

Ware, R W, Clark, D, Crossland, K, and Russel, R L, 1975:

"The Nerve Ring of the Nematode *Caenorhabditis elegans*: Sensory Input and Motor Output", Journal of Comparative Neurology, 162: 71 - 110 (1975)

White, J G, 1986: personal communication

White, J G, Albertson, D G, and Anness, M, 1978:

"Connectivity changes in a class of motoneurone during the development of a nematode",
Nature, 271: 764 - 766 (1978)

White, J G, Southgate, E, Thomson, J N, and Brenner, S, 1976:

"The Structure of the Ventral Nerve Cord of *Caenorhabditis elegans*",
Phil. Trans. R. Soc. London (Biol), 275: 327 - 348 (1976)

White, J G, Southgate, E, Thomson, J N, and Brenner, S, 1983:

"Factors That Determine Connectivity in the Nervous System of *Caenorhabditis elegans*",
Cold Spring Harbor Symp. Quant. Biol., 48: 633 - 640 (1983)

White, J G, Southgate, E, Thomson, J N, and Brenner, S, 1985:

"The Structure of the Nervous System of the Nematode *Caenorhabditis elegans*",
to appear in Phil. Trans. R. Soc. London (Biol), (unpub.)

Glossary of Terms

Acetylcholine

The most commonly encountered excitatory neurotransmitter.

Actuator

A system element which transforms a control signal into some definite action.

Amphid

A sensory organ, primarily sensitive to concentration of various chemical compounds, consisting of a multicellular pouch which is penetrated by the sensory endings of a number of neurons and has a pore opening to the outside.

Anion

A negative ion. Anions are so named because they are attracted to the anode in electrochemical experiments.

Aromatic compound

An organic compound containing one or more benzine rings.

Basis

A set of linearly independent vectors having the property that any vector in a given vector space can be expressed as a linear combination of the vectors in the set.

Basis vector

One of the vectors in a given basis.

Bifurcation

A two pronged fork in the trajectory of a system through its state space. As the state of a system advances along the single path leading to a bifurcation, it will choose one or the other of the paths leading away, as determined by very small random fluctuations in its state.

Boltzmann decision

The process by which a node in a Boltzmann machine determines its next state after sampling its input signal. The details of this process are internal to the node, and the only parameter of interest is the probability with which

the node chooses the active over the passive state.

Boltzmann distribution

The statistical distribution of states resulting when a group of one or more nodes repeatedly determine their states by making Boltzmann decisions.

Boltzmann's constant

One of the fundamental constants of nature. Boltzmann's constant expresses the amount of energy required to represent a single bit of information in a system having a temperature of one degree Kelvin.

Border matrix

A matrix formed by multiplying each element of the coupling matrix of a system by the second derivative of the transfer function of the node which receives input via that element. This matrix determines the conditions under which particular nodes are driven into or out of the flat parts of their transfer functions.

Border neurons

Those nodes in a system which are operating near the juncture between the flat and steep parts of their transfer functions. Such neurons contribute strongly to the bifurcations of the system dynamics.

Buccal

Of, or pertaining to, the mouth.

Campaniform

Literally, bell shaped. A hollow papilla.

Cation

A positive ion. Cations are so named because they are attracted to the cathode in electrochemical experiments.

Cephalic

Of, or pertaining to, the head.

Cervical

Of, or pertaining to, the neck.

Chemoreceptor

A neuron having structural specializations rendering it sensitive to the concentration of particular chemical compounds.

Cholinergic

A synapse which operates by the release of acetylcholine.

Collagen

A protein which forms the basis of connective tissue.

Commissure

A neural structure by which two other neural structures are joined together, or coupled.

Constraint

A coupling between two or more of the state variables of a system in such a manner that they are not completely free to vary independently.

Coupling

The degree to which the value of one state variable of a system is able to influence the value of another.

Coupling matrix

A matrix whose elements express the degree of coupling between the state variables of a system.

Critical neurons

Those nodes in a system which are operating near the centres of the steep parts of their transfer functions. Such neurons contribute strongly to the transfer operator of the system.

Cuticle

The outermost structural layer of the skin.

Degrees of freedom

The effective dimensionality of the state space of a system, taking all constraints into account. For example, a point constrained to move in a circular path has only one degree of freedom, even though its path is two dimensional.

Deirid

A sensory organ, primarily sensitive to vibration.

Diagonal operator

An operator which can be expressed as a diagonal matrix with respect to some basis.

Diagonalization

A procedure whereby a basis is found, with respect to which a given operator is diagonal.

Drive

A signal is said to drive a system when the signal is an argument of the transfer function of the system.

Eigenbasis

A basis of a vector space composed of the eigenvectors and generalized eigenvectors of a given operator defined on that space.

Eigenfunction

The set of all possible solutions of a differential equation form a vector space whose elements are functions. The eigenfunctions of the differential equation form a basis of this vector space in a manner analogous to the way in which the eigenvectors of a linear operator form a basis of the vector space on which the operator is defined.

Eigenvalue

The amount by which an eigenvector is scaled when operated on by a given operator. A linear operator is completely specified by giving an ordered list of its eigenvectors and associated eigenvalues.

Eigenvector

Any vector which is merely scaled, and not rotated, when operated upon by a given operator is called an eigenvector of that operator. The eigenvectors of a diagonalizable operator form a basis of the vector space on which the operator is defined. In effect, they define a set of preferred directions in this space.

Energy

The capacity to do work. That is, to exert a force and

thereby cause a change of state. Work is defined as the product of the force exerted and the amount by which the state changes.

Energy of interaction

The energy stored in a system as a result of an interaction between the elements of the system.

Entropy

The information neglected in a global description of a system. When the state of a system is characterized, for example, by its total internal energy, the information neglected expresses the number of different configurations of the system which result in each value of this energy.

Force

An interaction whereby one state variable of a system tends to cause another to change its value. Force is defined as the rate at which the energy of the system changes with this value.

Force constant

The constant of proportionality relating a change in the value of a state variable to the force required cause it.

Free energy

The sum of the mean value of the total internal energy of a system and the energy required to represent the neglected information, or entropy, of the system.

Gama-aminobutyric acid

The most commonly encountered inhibitory neurotransmitter.

Ganglion

A well defined substructure of the nervous system.

Gap junction

A very close apposition of specialized structures of the membranes of two adjacent neurons which permits electrical coupling between them, although this coupling is sometimes more effective in passing currents in one direction than the other.

Generalized eigenvector

A vector that, when operated on by a given operator, is transformed into one of the eigenvectors of the operator after a finite number of iterations.

High order constraint

A coupling between more than two of the state variables of a system in such a manner that they are not completely free to vary independently.

Homogeneous equation

A differential equation which lacks a forcing term.

Hydrostatic skeleton

The result of the incompressibility of water and the structure of the cuticle in certain invertebrates, which together impose constraints on the shape of the organism, which allows the muscles to act as though they had the benefit of skeletal attachment.

Hypodermis

A layer of cells located directly inside the cuticle.

Impulse response

The response of a system to a force consisting of a single impulse.

Information

A measure of the number of binary decisions required to distinguish between one state of a system and another. Thus the information content of a system in a given state is the binary logarithm of the probability of occurrence of that particular state.

Interaction matrix

A matrix formed by multiplying each element of the coupling matrix of a system by the first derivative of the transfer function of the node which receives input via that element. This matrix determines the effective strength of the interactions between the nodes.

Invariant subspace

A subspace of the vector space, on which an operator is defined, which has the property that any vector in this

subspace is transformed by the operator into another vector in the same subspace.

Jordan canonical form

An operator is said to be expressed in Jordan canonical form when its matrix is expressed relative to its eigenbasis. In this case, the eigenvalues of the operator appear in order along the main diagonal, the elements of the first superdiagonal are either one or zero, depending on the dimensionality of the invariant subspaces, and all other elements of the matrix are zero.

Klinokinesis

A method of orientation in which the organism measures a gradient by measuring the intensity of a stimulus at various points along its path.

Klinotaxis

A method of orientation in which the organism measures a gradient by measuring the intensity of a stimulus at the extrema of some regular motion.

Labial

Of, or pertaining to, the lips.

Lattice gas

A system consisting of a lattice of sites, each of which may, at any point in time, be in one of two states.

Lumbar

Of, or pertaining to, the back.

Markov random field

A set of random variables coupled only by binary constraints.

Mean firing rate

In the case of a fluctuating binary random variable, the reciprocal of the mean time interval between transitions from a given value to the other.

Mechanoreceptor

A neuron having structural specializations rendering it sensitive to mechanical deformation.

Membrane capacitance

The capacity of a cell membrane to store electrical charge when a potential difference exists between the inside and outside of the cell.

Membrane potential

The electrical potential difference existing, at any point in time, between the inside and the outside of a cell.

Membrane resistance

The resistance offered by a cell membrane to the flow of electrical current through the membrane.

Meniscus

The curvature of the surface of a fluid at the interface where a solid body pierces this surface.

Nematode

A class of roundworm comprised of aquatic, terrestrial and parasitic forms. Characterized by a cylindrical body, most nematodes are less than one millimetre in length, although some species reach lengths over one foot.

Neurohumour

A chemical compound, other than a neurotransmitter, such as a hormone, to which neurons respond in some way.

Neuromuscular junction

A structure, similar to a synapse, which couples a neuron to a muscle cell.

Neuropil

A structure, consisting of a large number of densely interconnected neurons, which performs a primary processing function.

Neurotransmitter

A chemical compound which is transported across a synapse in order to transmit neural signals from one neuron to another.

Nilpotent operator

An operator which transforms any vector into the zero vector after at most some fixed number of iterations. This number

is called the degree of the operator.

Noise

Random fluctuations in the value of some parameter.

Nucleotide

An organic compound consisting of a chain of groups, each consisting of a sugar, a phosphate and a base of the type found in nucleic acids.

Orthogonal

Literally, at right angles. Two vectors are orthogonal when their scalar product vanishes. In particular, two random signals are orthogonal when their correlation vanishes.

Orthogonal complement

The orthogonal complement of a subspace is that subspace consisting of all vectors whose inner product with any vector in the given subspace vanishes.

Orthokinesis

The tendency of an organism to become more active with the increase of intensity of some stimulus.

Papilla

A small, rounded structure projecting outward from the surrounding surface.

Pharynx

A muscular sac located between the mouth and the remainder of the digestive tract.

Phasmid

A sensory organ, primarily sensitive to concentration of various chemical compounds, similar to an amphid, but distinguished from it by the fact that its supporting structure is unicellular.

Projection matrix

A matrix representing a projection operator.

Projection operator

An operator which annihilates the component of a vector

lying in a particular subspace and leaves unchanged the component lying in the orthogonal complement of this subspace.

Proprioreceptor

A neuron having structural specializations rendering it sensitive to variations in the internal state of an organism.

Pseudocoelom

An annular, fluid filled cavity separating the inner body wall and its associated structures from the intestinal tract.

Rate code

A method of encoding information by causing the rate at which events, such as transitions, occur to vary as a function of the value of a parameter to be encoded.

Sensilla

A primitive sense organ located in the skin.

Seta

A bristle or stiff hair.

Sigmoid

Literally, S-shaped. A curve having a single inflection point separating regions in which the curvature is concave in opposite directions.

Somatic

Of, or pertaining to, the body.

Somatosensation

Sensation pertaining to bodily state of an organism.

Span

A set of vectors is said to span a vector space or a subspace thereof if any vector in the space or subspace can be expressed as a linear combination of the vectors in the set.

Spicule

A small, rigid spike.

Spin glass

A system composed of densely coupled magnetic dipoles, each of which can be oriented either parallel or antiparallel to an externally imposed magnetic field.

Stable state

A state of a system is said to be stable if the dynamics of the system are such that, if the system is in this state, any perturbation from it will be opposed.

Standard basis

The standard basis of a vector space is composed of a set of orthonormal vectors. That is, each vector has unit length and is orthogonal to all of the others.

State

The state of a system is given by a unique assignment of values to all of its state variables.

State space

A vector space whose standard basis is made up of the possibly scaled state variables of a given system.

State variable

A state variable is a parameter of a system which is relevant to the behaviour of the system.

Subspace

A subspace of a vector space is a vector space which is entirely contained within the original vector space.

Synapse

A structure whereby a neural signal can be transmitted unidirectionally from one neuron to another by the transport of a specific chemical compound across a narrow gap between the two neurons.

Taxis

Any method by which an organism orients itself in response to an external stimulus.

Temperature

The amount of energy required to represent the each bit of information neglected in describing a system in terms of the mean value of its internal energy.

Thermoreceptor

A neuron having structural specializations rendering it sensitive to temperature.

Transducer

A system element which transforms one type of energy into another. For example, an element which responds to light and produces electrical impulses.

Transfer function

The function describing the response of a system to any given stimulus.

Transfer operator

The operator which expresses the transfer function of a system.

Tropotaxis

A method of orientation in which the organism measures a gradient directly.

Turgor pressure

The pressure of the internal fluids of an organism.

Unit impulse

A force which imparts a unit impetus over a negligible interval of time, as in a hammer blow. The unit impulse is represented mathematically as a Dirac delta function of unit area.

INDEX

Acetylcholine 38
  Cholinergic 38, 116
Actuator 22-24, 109
Adaptation 11, 24, 34, 110-111, 113
Algorithm 3, 9-10, 14, 95
Amphid 34-35, 97-99, 101
Anion 97
Aromatic compound 97
Association 10-11, 29, 92, 116-118, 125, 144
  Context free association 117
  Context sensitive association 29, 88, 92-93, 96
    116-117, 124, 144
Asymmetric divergence 106, 119
Axiomatic system 1, 8-9
Background 17, 92, 100, 115, 124, 127, 145
Basis . . 21, 130, 132, 134, 151-152, 154, 156, 160, 162
  Basis vector 132
  Change of basis 14, 60, 130-131, 134-135, 152, 154, 160
  Coordinate system 131
  Coordinates 131, 149
  Eigenbasis 134, 151, 154-155, 160, 162
  Standard basis 134, 154, 156, 160, 162
Bayes' theorem 19-20
Behaviour . . . 2, 6, 8-12, 14-15, 21, 24, 25, 27-28, 34
    60-61, 64, 71, 91, 93, 95, 99, 103
    105-109, 119, 124, 145
Boltzmann machine . 11, 13, 15, 17-19, 21, 24, 28, 60, 86
    90-92, 94, 103, 104, 107, 110, 118
  Boltzmann decision 17, 19
  Boltzmann distribution 17, 19, 118
  Boltzmann's constant 16
  Link 12, 15, 16, 24, 27, 84, 94, 95, 103, 105
  Link weights 12, 16, 18, 24, 27, 94, 95, 105-108
  Node 12, 15-27, 61, 106-108, 118
  Non-equilibrium boltzmann machine . 11, 13, 15, 90, 94
Border matrix 93, 95, 121-122, 124, 144, 145
Border neuron 124-125, 130
Buccal 30, 102
Caenorhabditis elegans 11, 24, 29, 30, 32, 31, 33, 34, 86
    97, 98, 100, 103, 104, 116, 117
Campaniform 35
Cation 97
Cephalic 34-35, 102-103
Cervical 34, 39, 116
Chemoreceptor 34, 102
Code 17
  Encode 17, 19, 95, 109
  Rate code 17
Cognition 2, 5, 8, 10

Collagen 32
Commissure 35-37
Competition 8
Consciousness 2, 4
Constraint 23-24, 58, 79, 84, 88, 107, 109-110
  High order constraint 23, 107, 109
Context free association 117
Context sensitive association 29, 88, 92, 93, 96, 116, 117
  124, 144
Cooperation 8
Coordinate system 131
Coordinates 131, 149
Correlation . . 87, 94, 96, 98-99, 101, 102, 107, 114-115
Coupling . . . 12-14, 27, 29, 37, 54-55, 61-63, 66, 71-72
  76, 79-81, 84, 88-90, 92, 101-102, 104, 110
  120, 127, 146
  Connection . . 22, 34, 38, 39, 42-45, 47-52, 98, 100-102
  104-105
  Coupling coefficient 62, 71-72, 76
  Coupling matrix 12, 66, 92, 104, 110, 120, 127
  Decoupled 67, 71, 100, 151, 152, 156
  Gap junction 37-38, 101, 104, 116
  Link 12, 15-16, 24, 27, 84, 94-95, 103, 105
  Link weights 12, 16, 18, 24, 27, 94, 95, 97, 102
  105-108
  Neuromuscular junction 37-38, 46
  Synapse . . 11, 37-38, 61, 86, 103, 105, 109, 110, 113
  116, 122, 159
  Uncoupled 67
Critical neuron 124, 144
Curvature 55-57, 82, 149
Cuticle 30, 32-33, 35, 56
Damping 29, 58, 71, 115, 164
Declarative 6, 95
Deductive inference 3, 7, 21
Degree of freedom 7, 24, 32, 61, 62
Deirid 34, 36
Diagonal 131, 133, 154
  Diagonal matrix 133
Diagonalization 128, 130
Differential equation . . 13, 15, 28, 65-66, 68, 104, 129
  145, 151
  Equation of motion . . . 15, 27-28, 65, 68, 92, 131, 151
  Homogeneous equation 126, 128, 130
  Matrix differential equation 28, 129, 145
  Vector differential equation 66, 105, 128, 132, 151, 159
Dimensionality 56
Dissipation 94, 96
Distribution 5, 6, 17, 19, 99, 102, 105-106, 118-119, 125
  Boltzmann distribution 17, 19
  Gaussian distribution 119
  Posterior distribution 30, 35, 36-37, 106
  Prior distribution 24, 106, 108, 109

Domain of discourse 1, 7, 24, 109
Drive . 7, 11, 16, 18-19, 22, 24, 28, 29, 37, 54, 55, 61
71, 81, 87, 89, 90, 97, 100, 109, 113, 117, 126
Undriven 126
Dynamical 8, 12, 17, 29, 93-94, 125
Dynamics 8-10, 14, 24, 29, 53, 54, 61, 71-72, 86-87, 90-92
94-95, 104, 110, 117, 124-125, 131, 145
Stochastic dynamics . 9-10, 86-87, 90, 92, 95, 134, 151
154-155, 160, 162
Eigenfunction 68, 134-143, 138-143
Eigenvalue . . . 131-132, 136-137, 144, 153, 156, 158, 159
Eigenvector . 68, 131-132, 135-136, 144, 153-154, 159-160
Generalized eigenvector 68, 132, 135, 144
Endocrine 34, 109
Energy . . . 5-6, 9-10, 16, 18, 19, 21-22, 25, 87, 94, 96
Free energy 18
Entropy 7, 10-11, 18-19, 21
Environment 21-23, 33, 84, 100-101, 104-107, 109
Epistemology 5
Equilibrium . . . 11-13, 15, 18-19, 21-22, 25, 27-28, 60
80-81, 90, 92, 94, 104, 119-120, 125, 127, 158
Estimation 20, 106, 108
Evidence 19-21
Self evidence 7
Evolution . . 5, 18, 68, 70, 72, 74-79, 92, 96, 153, 156
Excitation . . 38, 54, 57-58, 62, 64-66, 68-71, 74, 76-79
89, 156, 163
Expectation 19, 28, 104, 119
Expected value 104, 114, 127, 152
Exponential 25, 71, 126, 128, 158, 164
Feature extraction . . 7, 11, 23, 31-32, 72, 84, 92-93, 98
101-102, 117
Filter 13, 30, 37, 91-92, 94, 117
Linear filter 117
Novelty filter 91, 94
Skeleton filter 13, 92, 94, 144
Flexion 32, 38, 54-55, 60, 64, 80, 82, 87, 150
Fluctuation . . . 9, 15, 18, 23, 27, 29, 71, 113-114, 152
Force 5-6, 13-14, 18, 25-26, 32, 33, 55, 58, 59
61, 87, 96, 101, 150
Force constant 58
Forcing function . . . 90, 98-99, 126-127, 137, 146, 148
Fourier series 76
Free energy 18
Frequency 72, 82
Friction 33, 58-59, 65
Function . . . 4, 16-17, 19, 26, 27, 34, 54, 56, 64-65, 84
89, 94, 103, 111, 118, 124, 126-127, 129, 134
137-139, 144-145, 149
Eigenfunction 68, 134-143, 138-143
Forcing function 126-127
Transfer function . . 17, 27, 64, 84, 89, 94, 111, 118
124, 137

Gama-aminobutyric acid 38
Ganglion 35-37, 116
Gap junction 37-38, 101, 104, 116
Gating 14, 60, 64, 72, 79, 82, 87-90
Gaussian 119
Generalized eigenvector 68, 132, 135, 144
Habituation 99-100, 105
High order constraint 23, 107, 109
Homogeneous equation 126, 128, 130
Hydrostatic skeleton 32
Hypodermis 30, 32
Hypothesis 5, 8, 19-21
Impulse response 129-130
Inductive inference 3, 4, 7, 21
Inference 3, 4, 7, 11, 19-21
  Bayesian inference 19, 20
  Deductive inference 3, 7, 21
  Inductive inference 3, 4, 7, 21
  Maximum entropy inference 7, 11
  Statistical inference 19
Information 2, 4-10, 12, 14, 16-20, 22, 24, 61-62
86, 94-96, 98, 102-103, 105-109, 119, 122
  Entropy 7, 10-11, 18-19, 21
  Information processing . . 2, 4, 6-7, 10, 36-38, 40, 54
62, 91, 93-97, 100-101, 110, 116
  Postprocessing . 11-14, 24-25, 29, 37, 39, 53, 60, 86-88
90-91, 93, 97, 125
  Preprocessing . . . 11-12, 23, 86, 91, 97, 103-104, 110
112-114, 116, 117, 119, 122, 124, 127
136, 144
Information theory 5, 105
Input 11-12, 15-18, 20, 22-29, 33-36, 60-61
84, 89, 91, 94-95, 100-101, 104-105, 110-113
117, 119, 122, 124, 125-127, 129, 130
136-137, 144
Integrating factor 128, 155, 162
Intelligence 1-4, 8, 11, 24
Interaction . 6, 8, 12, 15, 16, 21-22, 27, 54, 55, 62, 66
71, 80, 88, 93, 95, 112-113, 117, 120, 122
124, 144, 145
  Energy of interaction 16
Interaction matrix . . 27, 66, 71, 93, 112, 120, 122, 124
144, 145
Interneuron 11, 22, 38, 116, 124, 145
Invariant subspace 130-132, 137, 144, 159
Jordan canonical form 131, 153-154, 162
Klinokinesis 98-99
Klinotaxis 97
Labial 34-35, 101-102
Lattice gas 17
Learning 23, 107, 109, 111
Linear 28, 69-70, 74, 75, 92, 117, 130, 132, 134
140, 142, 151, 157

Linear filter 117
Linear system 28, 92, 117
Linearize 28, 70-72, 93, 137, 151
Link 12, 15-16, 24, 27, 84, 94-95, 103, 105
  Link weights 12, 16, 18, 24, 27, 94-95, 105-108
Logic 1
Lossy integrator 24, 61, 125
Lumbar 36-37
Markov random field 107
Matrix 12, 16, 27-28, 32, 66-67, 71, 92-93, 95
104-105, 110-113, 120-122, 124, 127-135, 144
145, 151, 154, 156, 160
  Border matrix 93, 95, 121-122, 124, 144, 145
  Coupling matrix 12, 66, 92, 104, 110, 120, 127
  Interaction matrix . 27, 66, 71, 93, 112, 120, 122, 124
144-145
  Jordan canonical form 131, 153-154, 162
  Matrix differential equation 28, 129, 145
  Matrix exponential 128
  Projection matrix 112-113
Maximum entropy inference 7, 11
Mean 17, 18, 26-27, 54, 65, 67-70, 76, 79, 89
92, 94, 108, 130, 151-153, 156
Mean firing rate 17
Mechanoreceptor 34
Membrane capacitance 25, 125
Membrane potential . 25, 28, 55, 64, 72, 79, 104, 118-120
122, 126-127, 145
Membrane resistance 25, 125
Meniscus 33, 150
Metric 106
Motor neuron 11, 22, 35, 37-39, 47-52, 72, 97
Musculature . . . 11, 22, 29-30, 32, 33, 36, 37-41, 54-59
61-62, 64-66, 69, 72, 74, 76-79, 82, 84, 87
90, 102, 116, 163
Nematode 11, 30, 34, 103
Nerve cord 35-37, 39, 44-45, 60-62, 79, 84, 87-90
92, 97, 101, 104, 105, 112-117, 119, 121-122
124-125, 128, 136-137, 144-145, 156
Nerve ring 12-13, 34-38, 54, 60, 62, 64, 66, 68
70-72, 79, 82, 84
Nervous system . 7, 10-12, 21-24, 28-30, 33-35, 40-41, 57
84, 86-89, 91, 96-98, 104, 109, 110, 116
Network 15, 90, 95, 144
  Neural network 90, 95
  Skelton network 144
Neural network 90, 95
Neurohumour 33-34
Neuromuscular junction 37-38, 46, 97, 102
Neuron 11, 15, 17, 22, 25, 30, 32, 34-35, 37-40
47-52, 54, 64-66, 69, 70, 72, 74-79, 88-90
92, 94, 97-104, 110-111, 113-114, 116-118
122, 124-127, 130, 137, 144, 158, 163

Border neuron 124-125, 130
Critical neuron 124, 144
Interneuron 11, 22, 38, 116, 124, 145
Motor neuron 11, 22, 35, 37-39, 47-52, 72, 97
Sensory neuron 11, 22, 35, 37, 100-104, 111
Neuropil 35, 101, 116
Neurotransmitter 38, 110, 113
Acetylcholine 38
Gama-aminobutyric acid 38
Nilpotent operator 130-131
Node 12, 15-27, 61, 106-108, 118
Noise . . 7, 9-10, 12, 14, 18, 27, 29, 37, 61, 66, 89-90
94-95, 103, 113-114, 119
Thermal noise 7
Non-equilibrium 11, 13, 15, 18, 25, 90, 94, 119
Nonlinear 17, 28, 69, 70, 76, 92-93, 100, 110
117, 141, 143
Nonlinear system 28, 76, 92-93
Novelty filter 91, 94
Nucleotides 97
Operator 23-24, 27, 91, 112-113, 121, 130, 132
Projection operator 113
Transfer operator 27, 112-113, 121
Orientation 55-56, 99
Orthogonal 21, 105, 112
Orthogonal complement 105, 112
Orthokinesis 34
Oscillation 14, 29, 55, 60, 64-65, 71-72, 79, 82
87-90, 137, 146, 149
Oscillator 14, 54-55, 88
Output 15-17, 21-23, 26, 36, 54, 60, 70, 72, 89
100, 104-105, 110-114, 117, 124, 125
Papilla 34, 103
Perturbation 92, 115, 124, 127, 136, 144
Pharynx 30, 32, 35-36, 116
Phase 54-55, 79-82, 88, 146, 148-149, 164
Antiphase 65
Phasmid 35-36, 97
Posterior distribution 30, 35-37, 106
Postprocessing . . 11-14, 24-25, 29, 37, 39, 53, 60, 86-88
90-91, 93, 97, 125
Posture 33, 56, 64, 71, 149, 163
Preprocessing 11-12, 23, 86, 91, 97, 103-104, 110
112-114, 116-117, 119, 122, 124, 127
136, 144
Prior distribution 24, 106, 108, 109
Probability 17, 19-21, 26, 105-109
Conditional probability 105, 107
Problem 1, 4-7, 24, 56, 85, 106, 108-109
Procedural 6, 95
Processor 96
Projection 35-38, 91, 102, 105, 112-113, 135
Projection matrix 112-113

Projection operator 113
Proprioreceptor 33-35, 101
Pseudocoelom 32, 36
Random 9, 23, 26-27, 71, 98-100, 107, 113-114, 152
Markov random field 107
Rate code 17
Reasoning 3, 21, 25, 145, 163
Relevance 8, 10
Representation 3-5, 7, 16, 21, 60, 64, 66-67, 70, 107, 119
Response . . . 6, 15, 22-25, 34, 55, 64, 69, 72, 74-79, 87
89, 91, 95, 97, 100, 102, 104, 105, 109-112
118, 122, 124-125, 127, 129-130, 136-137, 144
Self organisation 6-8, 23-24, 92, 94, 110
Sensilla 30, 34-35, 97, 101-103
Sensory neuron 11, 22, 35, 37, 100-104, 111
Setae 103
Sigmoid 17, 27, 64, 84, 89, 94, 96, 111, 118, 124
137-139
Signal . . . 7, 9, 12, 15, 20-22, 24-28, 55, 60-61, 64, 67
71-72, 76, 79-81, 84, 87-90, 92-96, 98-101
104-105, 110-114, 117, 119, 125, 129, 145
151-152, 163
Sinusoid 72, 80, 137, 146, 148-149
Skeleton filter 13, 92, 94, 144
Somatic . . 13, 29, 33, 36, 39-41, 65-66, 82, 84, 87, 97
Somatosensory 35
Space 12, 18, 25, 28, 87, 93, 95, 105-106, 117
119, 130, 145, 146
State space . . . 12, 18, 25, 28, 87, 93, 95, 117, 119
130, 145
Subspace 112, 131-132, 137, 144
Spicule 30
Spin glass 19
Standard basis 134, 154, 156, 160, 162
State 12, 15-18, 23, 25, 28, 60-62, 64-65, 72, 76
82, 87-89, 93, 95, 104-105, 111, 113, 114, 117
119-120, 124-125, 130, 145, 151
Stable state 18, 87, 93, 95, 117, 119, 130, 145
State space 12, 18, 25, 28
State variable 61-62, 64, 76, 88, 125, 151
Stationary 26, 64, 109, 115, 119, 127, 145
Statistical inference 19
Statistics 26-27, 93, 109
Statistical 9, 19, 25, 28, 84, 92
Stimulus 23, 101-102, 125
Stochastic 9-11, 15, 17, 23, 26-27, 64, 84, 86-87
90-92, 94-96, 100, 109, 115, 152
Stochastic dynamics 9-10, 86-87, 90, 92, 95
Stochastic processor 96
Stochastic system 15, 23
Storage 94, 125
Structure . . 2, 5-7, 10-11, 28-30, 32, 39, 53-54, 56, 61
67, 84, 91, 95, 100, 102, 106, 110, 130, 133

144-145
Subspace 112, 131-132, 137, 144
Invariant subspace 130-132, 137, 144, 159
Superposition 64, 95, 100, 128, 137
Symbol 2-5, 9, 26, 95, 104
Abstract symbol 2, 3, 5, 9
Symbol manipulation 2, 3, 5, 95
Synapse . . . 11, 37, 38, 61, 86, 103, 105, 109, 110, 113
116, 122, 159
Synaptic plasticity 109
Synchronization 60, 87
System . . . 1-16, 18, 19, 21-30, 33-35, 37, 39-41, 53-54
57, 59-61, 67-68, 71-72, 76, 79, 84, 85-99
101-114, 116-120, 122, 124-125, 127-131, 136
144-145, 156, 158
Axiomatic system 1, 8-9
Linear system 28, 92, 117
Nervous system 7, 10-12, 21-24, 28-30, 33-35, 40-41, 57
84, 86-89, 91, 96-98, 104, 109-110, 116
Nonlinear system 28, 76, 92-93
Self organising system 8
Stochastic system 15, 23
Subsystem 32, 62, 64-65, 67, 68, 82, 87, 151, 153, 155
Taxis 34, 98
Taylor series 120, 122
Temperature . . . 6, 9, 12, 15-18, 25, 27, 34, 60-61, 92
94-95, 97, 100, 118-119
Equivalent noise temperature . . 9, 12, 27, 61, 94, 119
Thermal noise 7
Thermodynamics 5, 9, 21, 60, 92, 94, 125
Thermoreceptor 34
Time constant . 24-25, 61, 65-66, 70, 72, 79, 82, 125-126
Trajectory 150
Transducer 104
Transfer function . . . 17, 27, 64, 84, 89, 94, 111, 118
124, 137
Transfer operator 27, 112-113, 121
Transform 14, 60, 130-131, 134-135, 152, 154, 160
Transition 26, 28-29, 92-93, 95, 117-118, 124
Tropotaxis 97
Turgor pressure 32
Uncertainty 3, 4, 7, 12, 21, 94
Variance 26, 89, 92, 118-119, 122
Variation 27, 65, 67, 70, 76, 79, 89, 92, 95, 110
120-122, 144, 146, 151-153, 156, 163
Vector 16, 27, 66, 104-105, 128-130, 132-133, 135
151, 154, 160, 163
Basis vector 132
Eigenvector 68, 131-132, 135, 136, 144, 153-154
159-160
Generalized eigenvector 68, 132, 135, 144
Vector differential equation 66, 151
Velocity 82, 146, 148, 150

Wave 32-33, 36, 38, 54-55, 60, 64, 80, 82, 87-88
146-150
Standing wave 64
Travelling wave 32, 55, 64, 88, 146-147, 150
Waveform 72, 76, 90, 94, 146, 149
Weight 12, 16, 18, 24, 27, 94-95, 105-108