Graduate Theses, Dissertations, and Problem Reports

1999

Identification and control of dynamical systems

Eugene Robert Mihaliuk
West Virginia University

Follow this and additional works at: https://researchrepository.wvu.edu/etd

Identification and Control
of Dynamical Systems

Eugene Mihaliuk

DISSERTATION

Submitted to the Eberly College of Arts and Sciences of
West Virginia University
in Partial Fulfillment of the Requirements for
the Degree of Doctor of Philosophy

Kenneth Showalter, Ph.D., Chairman
Paul Jagodzinsky, Ph.D
Reuben Symoi, Ph.D.
Charles Jaffé, Ph.D.
Mohindar Seehra, Ph.D

Morgantown, West Virginia
1999

Abstract

Identification and Control
of Dynamical Systems

Eugene Mihaliuk

Practical methods, based upon linear systems theory, are explored for applications to nonlinear phenomena and are extended to a larger class of problems. An algorithm for stabilizing, characterizing, and tracking unstable steady states and periodic orbits in multidimensional dynamical systems is developed and applied to stabilize and characterize an unstable four-cell flame front of the Kuramoto-Sivashinsky equation with six unstable degrees of freedom. A new method is presented for probing chemical reaction mechanisms experimentally with perturbations and measurements of the response. Time series analysis and the methods of linear control theory are used to determine the Jacobian matrix of a reaction at a stable stationary state subjected to random perturbations. The method is demonstrated with time series of a model system, and its performance in the presence of noise is examined. A new theory based on the construction of a multitude of linear models, each serving to represent one small region of the phase space, is presented together. Details of its implementation are presented in predicting chaotic Kuramoto-Sivashinsky wave fronts, demonstrating how it overcomes some of the problems associated with high dimensionality phase spaces. Motivated by the relationship between nonlinear prediction methods and the capabilities of neural systems, we demonstrate the possible role of nonlinear phenomena in the morphogenesis of neural tracts.

Acknowledgements

Contents

List of Figures

viii

Chapter 1

Introduction

Significant progress has been made over the past two decades in developing a theoretical understanding of nonlinear phenomena;[1,2] however, the gap between theory and practice is still considerable. We seek to narrow this gap by exploring methods based upon linear systems theory. We demonstrate how linear methods can be applied to complex nonlinear phenomena and, equally important, determine when these methods fail. We seek to improve the range of applicability by carefully extending the linear methods so that they may be used in the nonlinear regime. We draw upon existing theories of linear identification and control, striving to preserve the efficiency and simplicity pertinent to the linear approach.

We begin in Chapter 2 with the application of linear control theory[3] for the stabilization of a strongly nonlinear system in the local vicinity of a stationary steady state. An algorithm for stabilizing, characterizing and tracking unstable steady states and periodic orbits in multidimensional dynamical systems is presented, which provides an explicit connection between phase space approaches and time series analysis. The algorithm is applied to the stabilization and characterization of an unstable four-cell flame front of the Kuramoto-Sivashinsky equation,[4] with six unstable degrees of freedom.

We further explore the connection between time series and phase space of linearized dynamical systems in Chapter 3, where a method is presented for probing chemical reaction mechanisms experimentally with perturbations and measurements of the response. Time series analysis and the methods of linear control theory are used to

determine the Jacobian matrix of a reaction at a stable stationary state subjected to random perturbations. The method is demonstrated with time series of a model system, and its performance in the presense of noise is examined. We show that the Jacobian matrix of a chemical system in a stable stationary state can be determined from the time series of a single observable variable by applying perturbations to the relevant chemical species.

In Chapter 4 we present an extension of linear methods for prediction in spatial-temporal systems in a general domain, as opposed to a nearly linear vicinity of the steady state in the previous work. The attempt to extend the range of applicability, while preserving the simplicity of linear models leads to the introduction of the locally-linear method. This approach introduces a multitude of linear models, each representing one small region of the phase space. Thus, at any given moment in time, a single linear model is used, which preserves its amenability to the analysis by methods generally applicable to linear models. At the same time, as the system traverses the phase space, the model used is constantly changing, allowing for unrestricted topological complexity of the phase space, regardless of the effective dimensionality of the local model. The approach falls into the general class of State Dependent Models (SDM).[5] However, previously published examples of SDM have been limited to low-dimensional cases, whereas we have been able to demonstrate the accurate prediction of the dynamics of a spatially extended chaotic system. This has been possible due to an efficient implementation of the multidimensional prediction method, which is based upon a dynamically restructured mesh maintaining the structure of the flows in phase space. Since the efficiency of the high-dimensional implementation is of singular importance for the practical applicability of the method, we consider the development of this mesh-based method to be a significant improvement over previously known techniques. The performance of the method is verified by its application to the temporal and spatial prediction of the Kuramoto-Sivashinsky wave front.

With the goal to explore the similarities between the incremental growth and re-organization of the phase-space mesh and information processing associated with self-organization in biological systems, we have studied the interactions between entrainment in coupled dynamical systems and mechanisms of learning in biological systems. There is a striking similarity between pattern-driven self-organization in

neural systems[6] and certain aspects of the behavior of coupled spatially extended chemical systems.[7] In Chapter 5, using a minimal model, we show that entrainment of the activity in circuits with sufficiently similar dynamical properties can occur on the basis of weak and diffuse coupling. We conjecture that a reorganization of Hebbian connections and a gradually increasing synchronization of activity in the source and target circuits give rise to a positive-feedback loop. Starting with an array of initially random connections, such systems converge to a stable and topologically consistent configuration. We present this organizing principle as a possible mechanism for fine-tuning of neural connections. We believe that the exploration of the similarities between nonlinear filtering and prediction methods and the capabilities of neural systems may provide clues to the biological mechanisms responsible for the separation of signals and other fundamental elements of the recognition processes.

1.1   References

[1] H. D. I. Abarbanel, R. Brown, and J. B. Kadtke, "Prediction in Chaotic Nonlinear Systems: Methods for Time Series with Broadband Fourier Spectra," *Phys. Rev. A* **41**, 1782-1807, (1990).

[2] E. Ott, C. Grebogi, and J. A. Yorke, "Controlling Chaos," *Phys. Rev. Lett.* **64**, 1196-1199 (1990).

[3] R. C. K. Lee, "Optimal Estimation, Identification, and Control," Res. Monograph **26**, (M.I.T. Press: Cambridge Massachusetts, 1964).

[4] Y. Kuramoto, "Instability and Turbulence of Wavefronts in Reaction-Diffusion Systems," *Progress of Theoretical Physics* **63**, 1885-1903 (1980).

[5] M. Pristley, "Non-linear and Non-stationary Time Series Analysis," (Academic Press 1988).

[6] L. C. Katz and C. J. Shatz, "Synaptic Activity and the Construction of Cortical Circuits," *Science* **274**, 1133-1138 (1996).

[7] D. Winston, M. Arora, J. Maselko, V. Gaspar and K. Showalter, "Cross-Membrane Coupling of Chemical Spatiotemporal Patterns," *Nature* **351**, 132-135 (1991).

Chapter 2

Stabilizing and Characterizing Unstable States in High Dimensional Systems from Time Series

2.1   Introduction

Major strides have been made over the past few years in controlling chaos in low dimensional systems.[1] Unstable periodic orbits have been stabilized in magnetoelastic strips,[2] electronic circuits,[3,4] laser systems[5,6] and chemical reactions,[7–9] and recent reports of stabilizing periodic rhythms in heart tissue[10] and inducing periodic and chaotic behavior in hippocampal brain tissue[11] have stimulated widespread interest. It is clear that new developments in controlling dynamical systems offer opportunities for potentially important practical applications.

Several theoretical approaches have been advanced for stabilizing periodic orbits in chaotic systems. The feedback method proposed by Ott, Grebogi and Yorke (OGY)[1,12] and the various modifications of this method have been the most popular. The OGY method is appealing because it is easily understood in terms of the system state in phase space. Stabilizing an unstable orbit simply involves perturbing the system such that the stable manifold of the orbit is targeted each return. Thus, the positions of the system state and fixed point are known (in a suitable Poincaré section), and the effect of the perturbation is explicitly defined. Other control meth-

ods, including the continuous feedback algorithm of Pyragas,[13,14] are described and compared to the OGY method by Alsing *et al.*[15]

In systems that can be described by effectively 1-D maps, the OGY method can be reduced to an algorithm that directly targets the fixed point rather than the stable manifold.[16,17] The reduced algorithm is attractive for experimental applications because it requires minimal computational effort.[3,5,7] It can also be easily modified to permit tracking unstable steady and periodic states through bifurcation sequences.[18,19] Unstable periodic orbits[4,20] and steady states[21–23] have also been tracked with techniques that minimize fluctuations around the targeted fixed point.

It is known that the simple map-based approach may fail, even when a system is low dimensional and governed by a 1-D map.[17] This arises when the fixed point is shifted away from the unstable manifold of the original attractor as the parameter is perturbed. In such cases, the perturbed system can no longer be described in terms of the shifted 1-D map, which causes the method to fail. Rollins, Parmananda and Sherard[24] have recently proposed a recursive algorithm that corrects for this effect. They added a linear recursive term to the map-based algorithm, following an earlier suggestion by Dressler and Nitsche[25] for modifying the OGY method when time-delay coordinates are used. This yields a one-variable, one-parameter method that allows stabilization in the otherwise pathological case when the fixed point is shifted away from the original unstable manifold.[8]

Stabilizing and tracking states with more than one unstable direction remains an important challenge. Such states are common in spatially extended systems, and techniques beyond those developed for low-dimensional systems will be required for controlling spatiotemporal chaos. Simple techniques may be successful in certain spatiotemporal systems, such as when the behavior is highly spatially correlated.[19] Spatiotemporal chaos has also been controlled in a convectively unstable system, where the stabilized behavior is swept into the surrounding regions.[26] Proportional feedback has been used to stabilize periodic behavior in a coupled map lattice by multiple pinnings at locally stabilized sites, where the density of sites is increased until ordered behavior is exhibited.[27] Auerbach *et al.*[28] have proposed a generalization of the OGY approach, applicable to systems with one unstable and many stable degrees of freedom. Romerias *et al.*[29] have developed an approach for stabilizing states with

multiple unstable directions and have applied this to a kicked double rotor model. It was necessary, however, to monitor all of the system variables for control.

In this chapter, we present a general method for stabilizing and characterizing states with many unstable degrees of freedom and possibly an infinite number of stable degrees of freedom. This generalization provides an explicit connection between the OGY and related phase space approaches and the linear control routines of classical single-input, single-output (SISO) systems.[30] The stabilization of high-dimensional unstable steady or periodic states requires only one system variable to be monitored and only one system parameter to be perturbed. The essential features of the approach are illustrated in Section 2.2 by considering a simple two-variable system, beginning with the case in which the system variables can be monitored directly and then generalizing this to a single experimental observable. A generalization of the approach to an n-dimensional system is described in Section 2.3. The algorithm is applied in Section 2.4 to stabilize and characterize an unstable four-cell flame front of the Kuramoto-Sivashinsky equation, which is found to have six unstable degrees of freedom. The method is also applied to stabilize and track a periodic orbit with two unstable directions. The advantages and limitations of the approach along with potential applications are described in Section 2.5.

2.2   Two-Variable System Illustration

A geometric description of the general stabilization method can be developed by considering its application to a simple two-variable system. The system behavior around the unstable steady state is described by two linearized equations. Discrete dynamics is assumed, reflecting an experimental setting in which the system is sampled and perturbed at a fixed rate. Oscillatory behavior in the vicinity of an unstable periodic orbit can be reduced to linear discrete-time equations by using a suitable Poincaré section.

For illustration purposes, we imagine that ξ_1 and ξ_2, the coordinates along the system eigenvectors, are monitored at regular time intervals to give the set of data pairs $(\xi_{1,i}, \xi_{2,i})$. If the ith point lies away from the fixed point $(\xi_{1,F}, \xi_{2,F})$ and if the

characteristic exponents describing the motion along the eigenvectors are λ_1 and λ_2, respectively, then the discrete-time equations of motion for the $i + 1$ point are

$$
\begin{aligned}
\xi_{1,i+1} - \xi_{1,F} &= \lambda_1(\xi_{1,i} - \xi_{1,F}), \\
\xi_{2,i+1} - \xi_{2,F} &= \lambda_2(\xi_{2,i} - \xi_{2,F}).
\end{aligned}
\tag{2.1}
$$

For convenience, we assume that the fixed point lies at the origin: $(\xi_{1,F}, \xi_{2,F}) = (0,0)$. If, however, we also impose a small perturbation on some parameter p, the position of the fixed point is shifted along some line in phase space by an amount proportional to the perturbation. Denoting this perturbation as p_{i+1}, the evolution equations now become

$$
\begin{aligned}
\xi_{1,i+1} &= \lambda_1 \xi_{1,i} + (1 - \lambda_1)\alpha_1 p_{i+1}, \\
\xi_{2,i+1} &= \lambda_2 \xi_{2,i} + (1 - \lambda_2)\alpha_2 p_{i+1}.
\end{aligned}
\tag{2.2}
$$

where the coefficients α_1 and α_2 are the projections of the shift vector $\partial \boldsymbol{\xi}_F / \partial p$ determining the change in the fixed point position along the corresponding eigenvectors with change in the parameter p.

If a second perturbation p_{i+2} is made at the next step, the second iteration will be given by

$$
\begin{aligned}
\xi_{1,i+2} &= \lambda_1^2 \xi_{1,i} + (1 - \lambda_1)\alpha_1(\lambda_1 p_{i+1} + p_{i+2}), \\
\xi_{2,i+2} &= \lambda_2^2 \xi_{2,i} + (1 - \lambda_2)\alpha_2(\lambda_2 p_{i+1} + p_{i+2}).
\end{aligned}
\tag{2.3}
$$

Note that the eigenvalues are assumed to be independent of the parameter perturbation, at least to leading order. From Eq. (2.3) it follows that $\xi_{1,i+2}$ and $\xi_{2,i+2}$ can be set equal to zero by an appropriate choice of the perturbations p_{i+1} and p_{i+2}, which are found as a solution of the linear system provided $\lambda_1 \neq \lambda_2$ and $\alpha_i \neq 0$. The requirement of the system eigenvalues to be nonequal and the parameter perturbation to displace the system along all of the unstable manifolds are the main conditions for achieving stabilization of multivariable systems using a single parameter.

8
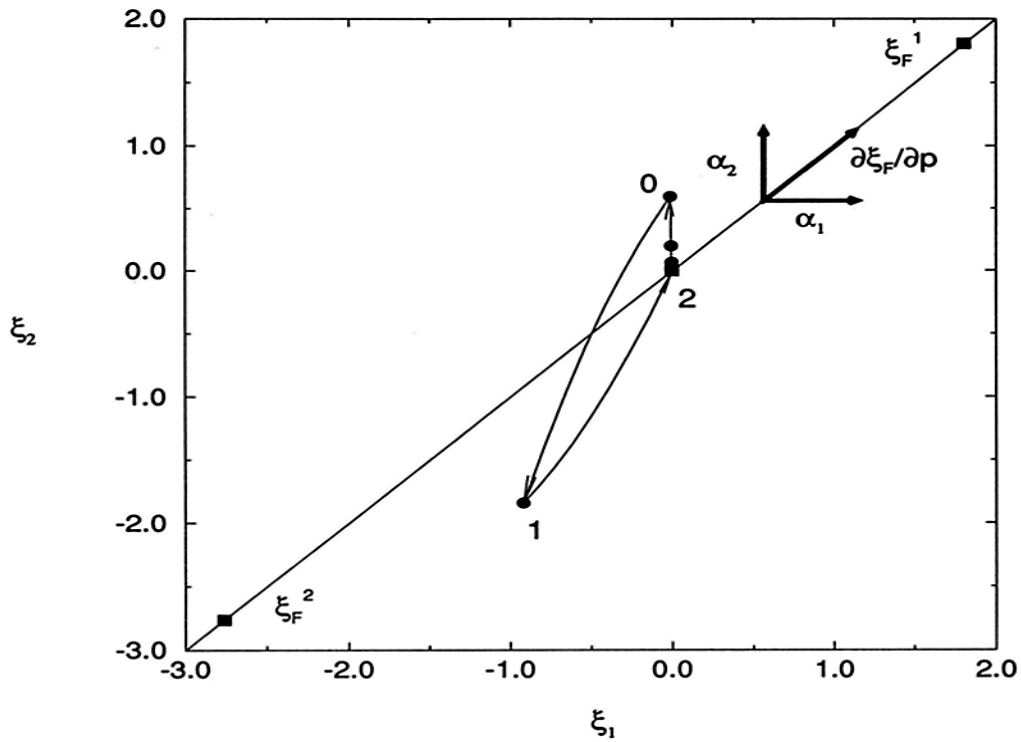
Figure 2.1: Geometrical description of control method. Application to a two-variable system with a fixed point characterized by $\lambda_1 = 1.5$ and $\lambda_2 = 3$. Two successive perturbations cause the system to evolve from point **0** to point **2**, corresponding to the fixed point of the unperturbed system.

Figure 2.1 shows an application with $\lambda_1 = 1.5$ and $\lambda_2 = 3$, where the system has evolved away from the unstable fixed point at the origin to the point marked **0**. In the first perturbation, the fixed point is moved along the shift vector $\partial \boldsymbol{\xi}_F / \partial p$ to $\boldsymbol{\xi}_F^1 = (\xi_{1,F}, \xi_{2,F})^1$. The system state evolves relative to the shifted fixed point according to the multipliers to the point marked **1**. In the second perturbation, the fixed point is moved again along the shift vector to $\boldsymbol{\xi}_F^2 = (\xi_{1,F}, \xi_{2,F})^2$. The system now evolves relative to this fixed point position to the point marked **2**, which corresponds to the targeted fixed point of the unperturbed system.

This example provides a graphical description of the targeting procedure; however, it differs from the operational procedure in that the sequence of n controlling perturbations must be determined in advance. In real-time applications of the algorithm, the value of the controlling perturbation is updated each step. For the two-variable example, the $i + 1$ perturbation is

$$p_{i+1} = k_1 \xi_{1,i} + k_2 \xi_{2,i}, \tag{2.4}$$

where the coefficients k_1 and k_2 are chosen to ensure that the fixed point is targeted on the second iteration. There are several possibilities for finding the appropriate k_1 and k_2 when the system coordinates or all independent system variables are known. The approach originally suggested by OGY targets the stable manifold of a state with one unstable direction.[12] If the fixed point has two unstable manifolds, it can still be targeted using Eq. (2.4) with two successive perturbations.[29] In the absence of noise, control equations (2.3) and (2.4) each produce the same sequence of two perturbations for targeting the fixed point.

In experimental settings, we commonly do not have access to the actual system variables, nor can we monitor n independent observable variables. Time-delay embedding techniques can be used to reconstruct the system state around a periodic orbit[12] provided that a correction for the shift of the Poincaré section is carried out.[25,28,29] Here we present a general approach for reconstructing the system state from the readings of one observable variable in the presence of perturbations. The approach is applicable to system dynamics on a Poincaré section or around a steady state.

In general, we monitor some observable x that is a linear combination of the system variables. For our two-variable illustration, we have

$$x_i = t_1\xi_{1,i} + t_2\xi_{2,i}. \tag{2.5}$$

The choice of variable x is largely unrestricted; however, the expected unstable behavior must be observable by monitoring x.

We will show that the system state vectors $\boldsymbol{\xi}$, scaled by projection coefficients t_i, can be reconstructed from successive readings of x and p. We rewrite Eqs. (2.2) for $i-1$ and then sum the first multiplied by $\lambda_2 t_1$ with second multiplied by $\lambda_1 t_2$, and with Eq. (2.5) for $i-1$ obtain

$$\lambda_1\lambda_2 x_{i-1} + [\lambda_2(1-\lambda_1)\alpha_1 t_1 + \lambda_1(1-\lambda_2)\alpha_2 t_2]p_i = \lambda_2 t_1\xi_{1,i} + \lambda_1 t_2\xi_{2,i}. \tag{2.6}$$

The system state $(t_1\xi_{1,i}, t_2\xi_{2,i})$ can therefore be found as a solution of Eqs. (2.5) and (2.6) from the x_i, x_{i-1} and p_i, provided that $\lambda_1 \neq \lambda_2$. The control equation (2.4) can then be written as

$$p_{i+1} = q_2 x_i + q_1 x_{i-1} + r_1 p_i, \tag{2.7}$$

where coefficients q_1, q_2 and r_1 can be expressed through the system parameters. It follows from Eq. (2.6) that the term involving x_{i-1} disappears from the control equation in the case when one of the manifolds is strongly attracting (for which $\lambda_2 \approx 0$). This corresponds to the recursive feedback algorithm of Rollins et al.[24] for the control of highly dissipative systems. It also follows from Eq. (2.6) that when the perturbation does not significantly shift the fixed point off the unstable manifold ($\lambda_2 \approx 0$ and $\alpha_2 \approx 0$) the simple map-based algorithm[16,17] is recovered.

We show in Section 2.3 that n successive readings of the observable x and $n-1$ values of the previous perturbations are required to control an n-dimensional system. We also show how the coefficients q_1, q_2 and r_1 of the control equation can be found, along with other system unknowns, from the time series of the experimental observable.

2.2.1 System identification and control

The control algorithm requires that the dynamical system be continuously interrogated by imposing random perturbations to a suitable parameter p at regular sampling

11

intervals. The time series obtained by recording some observable x then consists of a set of data pairs:

$$(x_1, p_1), (x_2, p_2), (x_3, p_3), \ldots, (x_i, p_i). \tag{2.8}$$

For a system with two independent variables, we need to record at least 7 data pairs (i.e., $3n + 1$) to allow identification of the system. Control can then begin with the parameter perturbation p_8.

The time series can be fitted to a recursive SISO (single-input, single-output) model[30] of the form

$$x_{i+1} = a_2 x_i + a_1 x_{i-1} + a_0 + b_2 p_{i+1} + b_1 p_i. \tag{2.9}$$

This form is also known as an ARX (auto-regressive with auxiliary input) equation.[31] The number of fitting coefficients a and b and their relation to the system parameters is discussed in Section 2.3. For the two-dimensional system $n = 2$ and there are five unknowns: a_0 - a_2, b_1 and b_2. Applying this approach to obtain the data pairs in (2.8) produces a set of five equations, which we can write explicitly as

$$
\begin{aligned}
x_3 &= a_2 x_2 + a_1 x_1 + a_0 + b_2 p_3 + b_1 p_2, \\
&\vdots \\
x_7 &= a_2 x_6 + a_1 x_5 + a_0 + b_2 p_7 + b_1 p_6,
\end{aligned}
\tag{2.10}
$$

where the coefficient a_0 is related to the fixed point x_F by

$$a_0 = (1 - a_2 - a_1)x_F. \tag{2.11}$$

We note that $a_0 = 0$ if $x_F = 0$, which we assume in the following treatment. [Alternatively, the optimal values of the coefficients could be determined more precisely from a longer data set using singular value decomposition, but here we proceed from this minimum basis set.]

The eigenvalues λ_1 and λ_2 governing the "autonomous" system, i.e., the system in the absence of perturbations, can be determined from the a_i coefficients. Motion along

the ith eigenvector occurs according to $x_n \propto \lambda_i^n$, and substituting the corresponding terms into (2.9) yields the characteristic equation for the system eigenvalues:

$$-\lambda^2 + a_2\lambda + a_1 = 0. \tag{2.12}$$

The eigenvalues of the "closed loop" system under control can be obtained by deriving the recursive model in a form that does not depend on the perturbations. We obtain the perturbation-independent equation by combining Eqs. (2.7) and (2.9) to express p_{i+1}:

$$p_{i+1} = \frac{r_1 x_{i+1} + (b_1 q_2 - r_1 a_2)x_i + (b_1 q_1 - r_1 a_1)x_{i-1}}{r_1 b_2 + b_1}. \tag{2.13}$$

Equation (2.13) for p_i and p_{i+1} can be substituted into Eq. (2.9), allowing x_{i+1} to be expressed as a linear combination of x_i, x_{i-1}, and x_{i-2}:

$$x_{i+1} = l_3 x_i + l_2 x_{i-1} + l_1 x_{i-2}, \tag{2.14}$$

where

$$\begin{aligned}
l_3 &= r_1 + a_2 + b_2 q_2, \\
l_2 &= a_1 - r_1 a_2 + b_1 q_2 + b_2 q_1, \\
l_1 &= b_1 q_1 - a_1 r_1.
\end{aligned} \tag{2.15}$$

The controlled system is described by Eq. (2.14) and is characterized by a total of three eigenvalues, λ_1, λ_2 and λ_3, which can be found as roots of the polynomial

$$-\lambda^3 + l_3\lambda^2 + l_2\lambda + l_1 = 0. \tag{2.16}$$

We achieve stabilization using the pole placement technique, i.e., by requiring these eigenvalues to adopt the target values λ_1^*, λ_2^* and λ_3^*. The roots of Eq. (2.16) will have the appropriate values of λ_i^* when

$$\begin{aligned}
l_1 &= \lambda_1^*\lambda_2^*\lambda_3^*, \\
l_2 &= -(\lambda_1^*\lambda_2^* + \lambda_2^*\lambda_3^* + \lambda_3^*\lambda_1^*), \\
l_3 &= \lambda_1^* + \lambda_2^* + \lambda_3^*.
\end{aligned} \tag{2.17}$$

13

There is some freedom in selecting the target values. In general, we require $|\lambda_i^*| < 1$ so the system converges toward the fixed point: the smaller the magnitude of the target eigenvalues, the faster will be the convergence. When all of the eigenvalues of the system under control are chosen to be zero (so-called "dead-beat" control) the system should converge to the steady state after n iterations. This, however, may involve the imposition of larger perturbations at the early stages of control. One approach is to leave the stable eigenvalues unchanged. This control results in targeting the corresponding stable manifolds rather than the fixed point and reduces the magnitude of the control perturbations. It is important to note that setting the eigenvalues close to unity is dangerous, since even small errors in the system identification can then make the system unstable. Also, if the system is high dimensional, errors in the system parameters that are inevitably carried over from the identification stage may become large. Selecting the optimal control law in the presence of system parameter errors is the subject of the H_∞ control approach.[32]

The various coefficients a_0 - a_2, b_1 - b_2 and l_1 - l_3 are now used to calculate the required perturbation to be imposed at the next time step:

$$p_8 = q_2 x_7 + q_1 x_6 + q_0 + r_1 p_7, \tag{2.18}$$

Here q_1, q_2 and r_1 are given by the solutions of Eqs. (2.15) and (2.17) and q_0 is assumed to be zero (which is equivalent to $x_F = 0$). This perturbation is applied at the seventh sampling time. The process is then repeated, with the appropriate perturbation p_9 being calculated from x_8, x_7, and p_8.

2.3   Generalization to n-Variable System

The control algorithm can be generalized to apply to a system of dimensionality n. The coefficients a_0 - a_n and b_1 - b_n are determined by fitting the recursive SISO model to at least $3n + 1$ data pairs collected from the interrogated system (i.e., the system subjected to random perturbations at each sampling time) and the target values selected for the eigenvalues of the controlled system.

If a sequence of n perturbations p_i is applied to an n-dimensional system, the

following equations can be written for n consecutive iterations of the state vector (i.e., the coordinates along the system eigenvectors):

$$
\begin{aligned}
\boldsymbol{\xi}_1 &= \boldsymbol{\xi}_1, \\
\boldsymbol{\xi}_2 &= \hat{\lambda}\boldsymbol{\xi}_1 + (\hat{I} - \hat{\lambda})p_2\boldsymbol{\alpha}, \\
\boldsymbol{\xi}_3 &= \hat{\lambda}^2\boldsymbol{\xi}_1 + (\hat{I} - \hat{\lambda})(p_2\hat{\lambda} + p_3)\boldsymbol{\alpha}, \\
&\vdots \\
\boldsymbol{\xi}_3 &= \hat{\lambda}^n\boldsymbol{\xi}_1 + (\hat{I} - \hat{\lambda})(p_2\hat{\lambda}^{n-1} + \cdots + p_n\hat{\lambda} + p_{n+1})\boldsymbol{\alpha},
\end{aligned}
\tag{2.19}
$$

where

$$
\hat{\lambda} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix}, \quad
\hat{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad
\boldsymbol{\alpha} = \frac{\partial \boldsymbol{\xi}_F}{\partial p}.
\tag{2.20}
$$

In the general case, the observable variable x is a linear combination of ξ_i,

$$
x = \mathbf{t} \cdot \boldsymbol{\xi}.
\tag{2.21}
$$

We can therefore rewrite the last equation of (2.19) as

$$
x_{n+1} = (\boldsymbol{\lambda}^n \hat{t}\boldsymbol{\xi}_1) + (\mathbf{p}_{n+1}\mathbf{L}(\hat{I} - \hat{\lambda})\hat{t}\boldsymbol{\alpha}),
\tag{2.22}
$$

where

$$
\mathbf{L} = \begin{bmatrix} \lambda_1^{n-1} & \cdots & \lambda_n^{n-1} \\ \vdots & & \vdots \\ \lambda_1 & \cdots & \lambda_n \\ 1 & \cdots & 1 \end{bmatrix}, \quad
\boldsymbol{\lambda}^n = \begin{bmatrix} \lambda_1^n \\ \vdots \\ \lambda_n^n \end{bmatrix},
\tag{2.23}
$$

and

$$
\hat{t} = \begin{bmatrix} t_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & t_n \end{bmatrix}, \quad
\mathbf{p}_{n+1} = \begin{bmatrix} p_2 \\ \vdots \\ p_{n+1} \end{bmatrix}.
\tag{2.24}
$$

15

We can further rearrange Eq. (2.22) to express x_{n+1} as a function of $\mathbf{x}_n = (x_1, \ldots, x_n)$ and $\mathbf{p}_{n+1} = (p_2, \ldots, p_{n+1})$:

$$x_{n+1} = (\mathbf{a} \cdot \mathbf{x}_n) + (\mathbf{a}\hat{t}(\hat{I} - \hat{\lambda})\mathbf{L}^T \mathbf{A}\mathbf{p}_{n+1}), \tag{2.25}$$

where

$$\mathbf{a} = \begin{bmatrix} 1 & \lambda_1 & \ldots & \lambda_1^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_n & \ldots & \lambda_n^{n-1} \end{bmatrix}^{-1} \boldsymbol{\lambda}^n, \quad \mathbf{A} = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ -a_n & 1 & \ldots & 0 \\ \vdots & & \ddots & \\ -a_2 & \ldots & -a_n & 1 \end{bmatrix}. \tag{2.26}$$

Equation (2.25) can be rewritten as

$$x_{i+1} = a_n x_i + a_{n-1} x_{i-1} + \cdots + a_1 x_{i-n+1} + a_0 + b_n p_i + \cdots + b_1 p_{i-n+1}. \tag{2.27}$$

Equations (2.25, 2.26, 2.27) provide the connection between the coefficients a_i, b_i and the phase space description of the system, as used in the OGY control algorithm. Specifically, the eigenvalues of the system are the roots of the polynomial

$$-\lambda^n + a_n \lambda^{n-1} + \cdots + a_2 \lambda + a_1 = 0, \tag{2.28}$$

and the b_i coefficients are linear combinations of the projections of the shift vector. Equation (2.25) is related to the Laplace transformation from the state space realization to the transfer function widely used in classical control theory.[30]

The $2n + 1$ unknown coefficients of Eq. (2.27) can be calculated from the time series of a single observable variable. If n is the dimensionality of the system, then $3n + 1$ successive readings of the variable and $3n + 1$ corresponding perturbations are required for the solution. Since n previous perturbations and observations are required to predict the future of the system according to Eq. (2.27), the control law should involve the same number of variables.

The control perturbation for the $i + 1$ step is calculated from the equation

$$p_{i+1} = q_n x_i + q_{n-1} x_{i-1} + \cdots + q_1 x_{i-n+1} + q_0 + r_{n-1} p_i + \cdots + r_1 p_{i-n+1}, \tag{2.29}$$

using the pole placement technique for recursive models.[30] The coefficients q_1 - q_n and r_1 - r_{n-1} are the solutions of the linear system

$$
\begin{bmatrix}
a_n & -1 & 0 & \dots & 0 & b_n & 0 & \dots & \dots & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & & \vdots \\
a_1 & & \ddots & \ddots & 0 & b_1 & & \ddots & \ddots & \vdots \\
0 & \ddots & & \ddots & -1 & 0 & \ddots & & \ddots & 0 \\
\vdots & \ddots & \ddots & & a_n & \vdots & \ddots & \ddots & & b_n \\
\vdots & & \ddots & \ddots & \vdots & \vdots & & \ddots & \ddots & \vdots \\
0 & \dots & \dots & 0 & a_1 & 0 & \dots & \dots & 0 & b_1
\end{bmatrix}
\begin{bmatrix}
1 \\
-r_{n-1} \\
\vdots \\
-r_1 \\
q_n \\
\vdots \\
q_1
\end{bmatrix}
=
\begin{bmatrix}
l_{2n-1} \\
\vdots \\
l_1
\end{bmatrix},
\quad (2.30)
$$

where the l_i are the $2n - 1$ coefficients that correspond to the target eigenvalues of the controlled system:

$$
-\lambda^{*2n-1} + l_{2n-1}\lambda^{*2n-2} + \cdots + l_2\lambda^* + l_1 = 0. \tag{2.31}
$$

We note that the overall dimensionality of the controlled system is increased by $(n-1)$.

In general, the dimensionality of the system will not be known in advance. Nor can the effective number of degrees of freedom always be deduced from the evolution of the autonomous system in the linear region of the unstable fixed point. The parameter perturbations used for control may shift the system onto stable manifolds not evident in the unperturbed case and reveal additional dimensions. On the other hand, the effective dimensionality can be determined by interrogating the system with the method outlined above. For spatiotemporal systems, most of the infinite number of modes will decay rapidly compared to the sampling period. We follow the suggestion by Auerbach *et al.*[28] of lumping all the highly attracting manifolds together as one. The effective dimensionality is therefore equal to the number of unstable and slowly attracting stable manifolds plus one. The method of determining n will be illustrated with reference to a particular example in the next section.

2.4 Stabilizing High Dimensional States of the Kuramoto-Sivashinsky Equation

The Kuramoto-Sivashinsky (KS) equation is one of the simplest nonlinear partial differential equations for modeling spatiotemporal chaos. It has been found to mimic the dynamical behavior of many different physical systems, but is most often used to model the spatiotemporal evolution of 2D flame fronts.[33] The governing equation for the contour of the front has the form

$$\frac{\partial \psi}{\partial t} + \left(\frac{\partial \psi}{\partial z}\right)^2 + \frac{\partial^2 \psi}{\partial z^2} + \frac{\partial^4 \psi}{\partial z^4} = 0. \tag{2.32}$$

We use the KS equation as an example of a multidimensional system that can be stabilized with the control algorithm. With a reaction zone width of $L = 35.0$, a symmetrical four - cell solution is found to be unstable and the system diverges away from this state to exhibit spatiotemporal chaos. Figure 2.2 shows the front as it moves away from the symmetrical state and the two quantities, y_1 and y_2, that serve as the "experimental" observables to monitor the spatiotemporal evolution of the profile. A two-dimensional projection of the phase portrait constructed from these observables is shown in Figure 2.3, where the system is evolving away from the unstable state. Although the evolution of the system is followed very close to the unstable state, it is clear that the behavior is high dimensional. Generally, the time series of only one observable variable provides enough information for control. Figure 2.4 shows the corresponding time series generated from the observable y_1, which is used as the monitored variable in the control algorithm. The points in Figures 2.3 and 2.4 show the sampling times of the monitored variable.

2.4.1 Determination of dimensionality

The evolution of the four-cell front away from the unstable symmetric solutionwhere a random perturbation is applied to a selected parameter each sampling intervalis shown in the identification part of Figure 2.5. The parameter chosen for perturbation is the gradient $\partial \psi / \partial z$ at z = 0. The unperturbed boundary condition, corresponding to the autonomous system in Figures 2.3 and 2.4, is $\partial \psi / \partial z = 0$ at this point. In

18
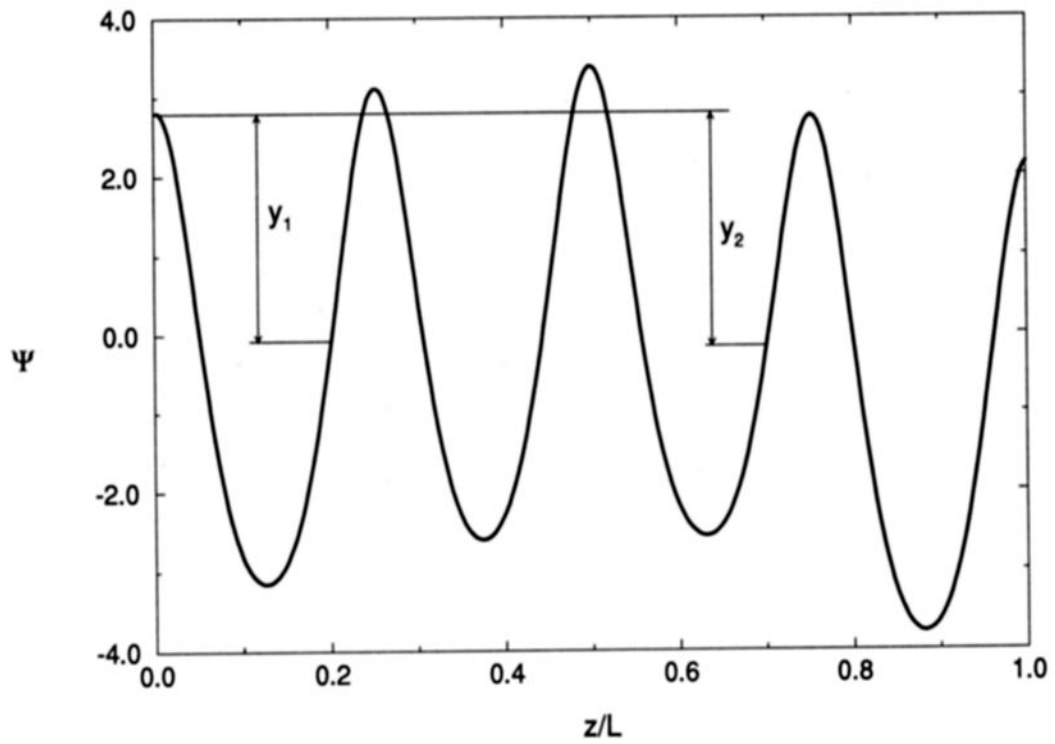
Figure 2.2: Four-cell solution of the KS equation for the system width $L = 35.0$. The symmetrical solution is unstable at this value of L, and this profile shows a snapshot in the early evolution away from the symmetrical state. The observables y_1 and y_2 are used to monitor the spatiotemporal evolution of the profile.
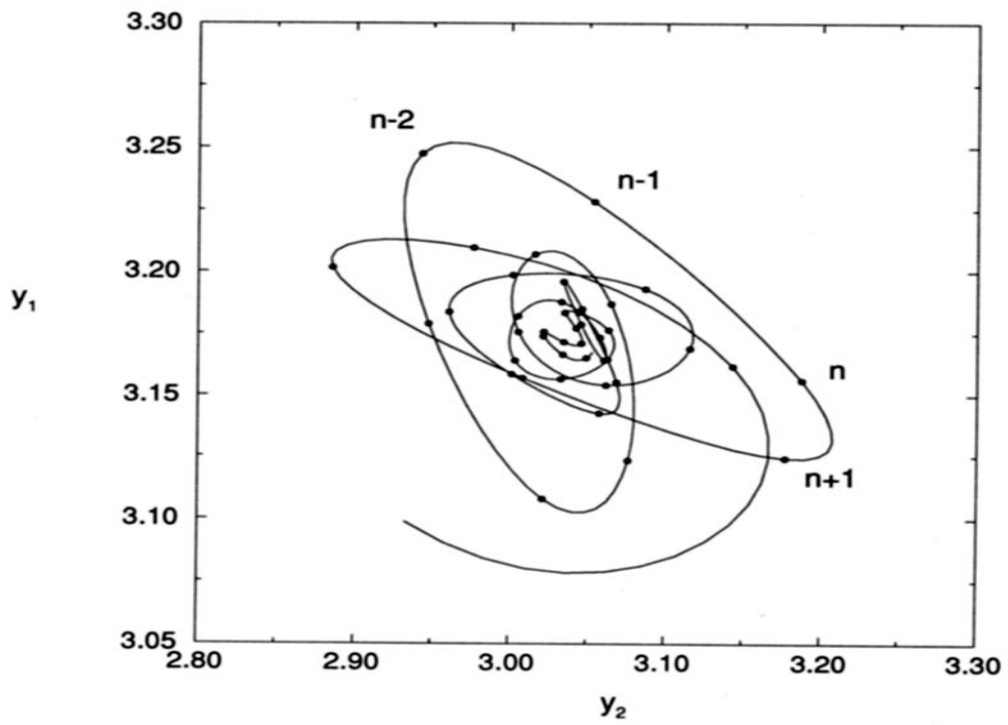
Figure 2.3: Phase portrait showing evolution of system away from symmetrical four-cell solution. Two-dimensional projection is constructed from the observables y_1 and y_2.
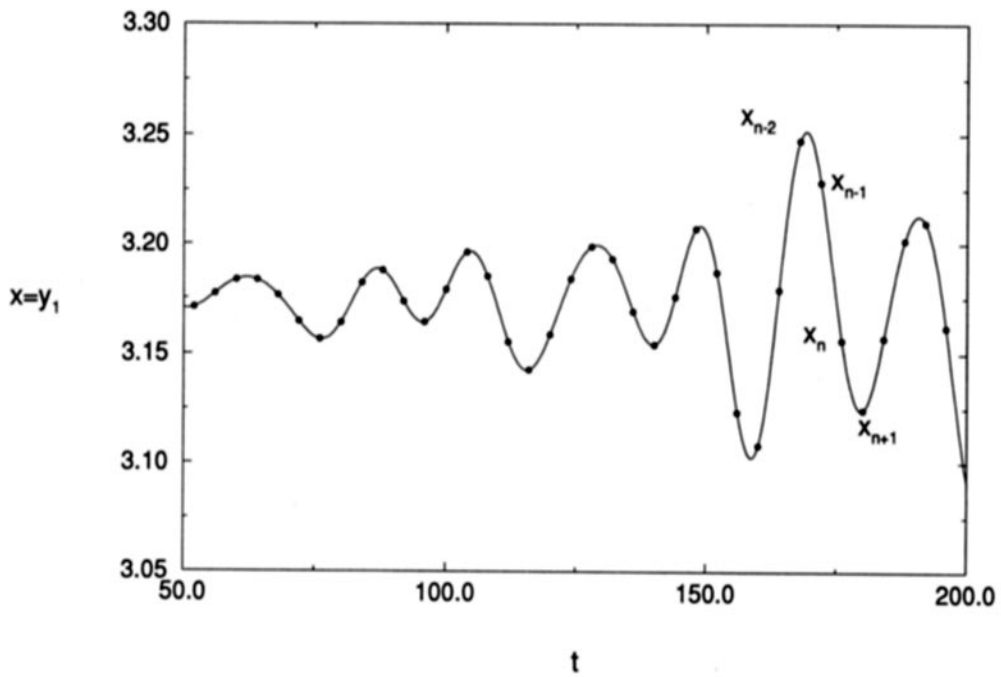
Figure 2.4: Time series showing observable y_1 used in identification and control as system evolves away from symmetrical four-cell solution.
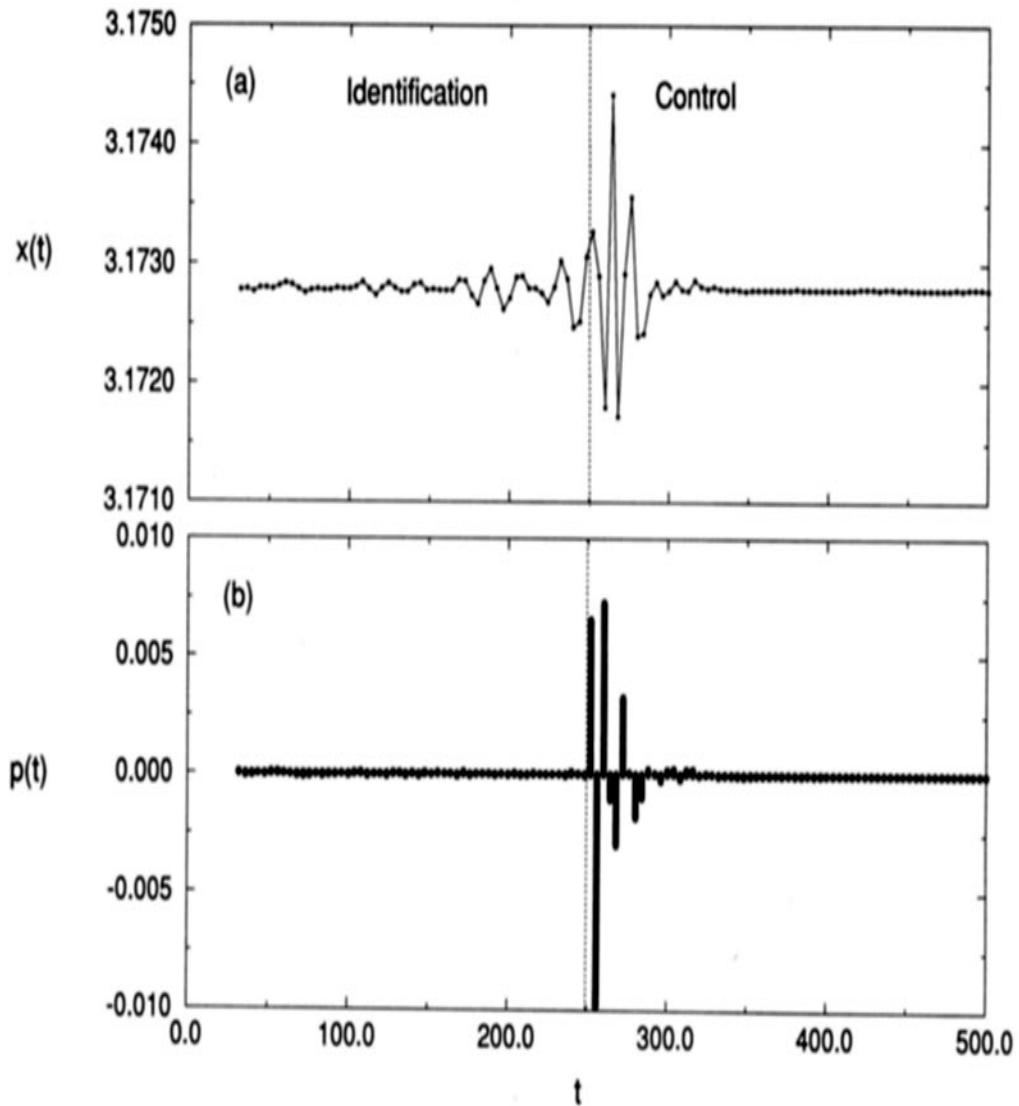
Figure 2.5: Identification and control of the unstable symmetrical four-cell solution of the Kuramoto-Sivashinsky equation. (a) Value of observable x_i during identification and control phases, and (b) value of the controlling parameter p_{i+1}.
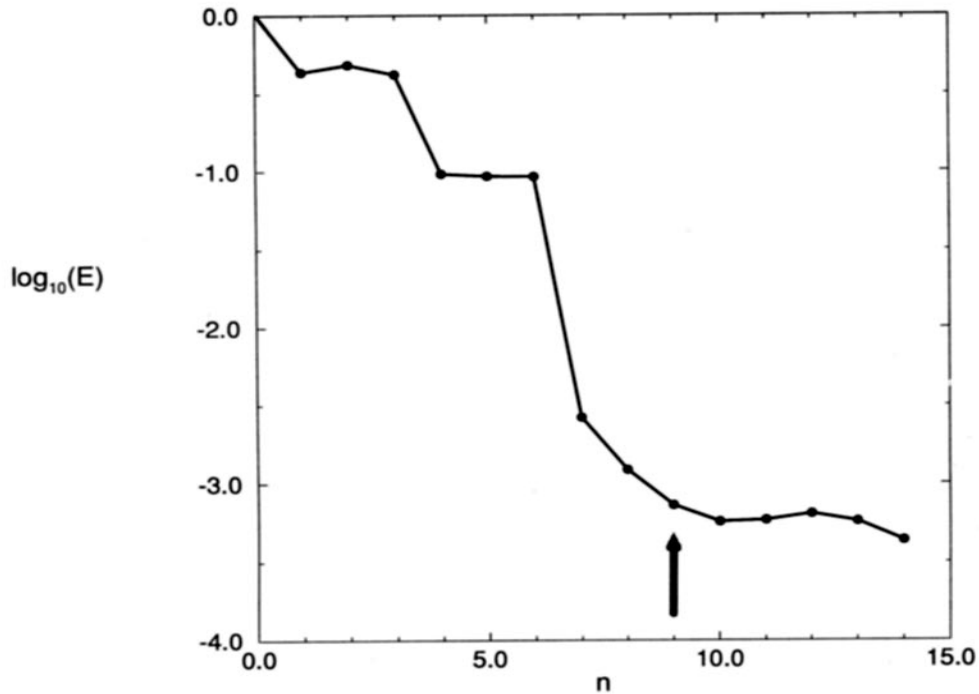
Figure 2.6: The dependance of the fitting error of Eq. (2.27) on n, showing plateau from $n = 9$. An effective dimensionality of $n = 9$ is used for stabilization of the symmetrical four-cell front shown in Figure 2.2. Error E is calculated with respect to standard deviation of predicted amplitude error relative to average amplitude of oscillation as shown in the identification stage of Figure 2.5(a).

general we utilize "mirror" boundary conditions, where the first and third derivatives in Eq. (2.32) are required to be zero at the boundaries.

The data pairs (x_i, p_i) from the perturbed system (up to $t = 250.0$ in Figure 2.5) are used to determine the dimension and the corresponding eigenvalues of the autonomous system. Error estimates for different choices of n are obtained by summing the error between the measurements and the optimized n-dimensional fit to Eq. (2.27) over the entire data set. Figure 2.6 shows the variation of this error as a function of n. For $n \geq 9$, there is no significant reduction in the fitting error on increasing the effective dimension, so we choose $n = 9$ for this system. It should be noted that the small error for convergence (typically 10^{-3} to 10^{-4}) suggests that it may be difficult to determine dimension from experimental data with this method due to the possibility of noise at

23

higher magnitudes than the convergence criterion. The eigenvalues of the unstable four-cell solution are then calculated as the roots of the equation

$$-\lambda^9 + a_9\lambda^8 + \cdots + a_2\lambda + a_1 = 0, \tag{2.33}$$

which are shown in Figure 2.7. There are six unstable eigenvalues ($\|\lambda\| > 1$) and three stable eigenvalues, the latter corresponding to modes excited by the parameter perturbations. The eigenvalue of smallest magnitude represents the lumping of an infinite number of stable modes that decay quickly compared to the sampling interval.

2.4.2 Stabilization of steady four-cell front

Once n and the coefficients a_i and b_i are determined, the algorithm is implemented in the control stage from $t = 250.0$. All the eigenvalues of the closed-loop system were chosen to be zero by setting $l_i = 0$ in Eq. (2.30). As indicated in Figure 2.5, the state is effectively stabilized after the first cycle of 9 perturbations. The control algorithm is applied continuously, with the values of the coefficients revised after each sampling and then used to calculate the next perturbation. In the present example, the magnitude of the control perturbations becomes very small after two cycles, or on the 19th iteration of the algorithm. We note that stabilization was also achieved with assumed system dimensionalities of $n = 10$ and 11. The higher order control laws are less desirable in experimental settings, however, due to their higher sensitivity to errors.

2.4.3 Stabilization of periodic two-cell front

The Kuramoto-Sivashinsky equation exhibits fronts with different numbers of cells on increasing the reaction zone width L. Each of these fronts loses its temporal stability through a bifurcation sequence that leads to chaotic behavior before the next front with more cells is established. We now examine the spatiotemporal behavior of a two-cell front in order to apply the control algorithm to a periodic orbit with more than one unstable direction. Specifically, we will stabilize and track a period-1 limit cycle through a secondary Hopf bifurcation, where the orbit is characterized by two
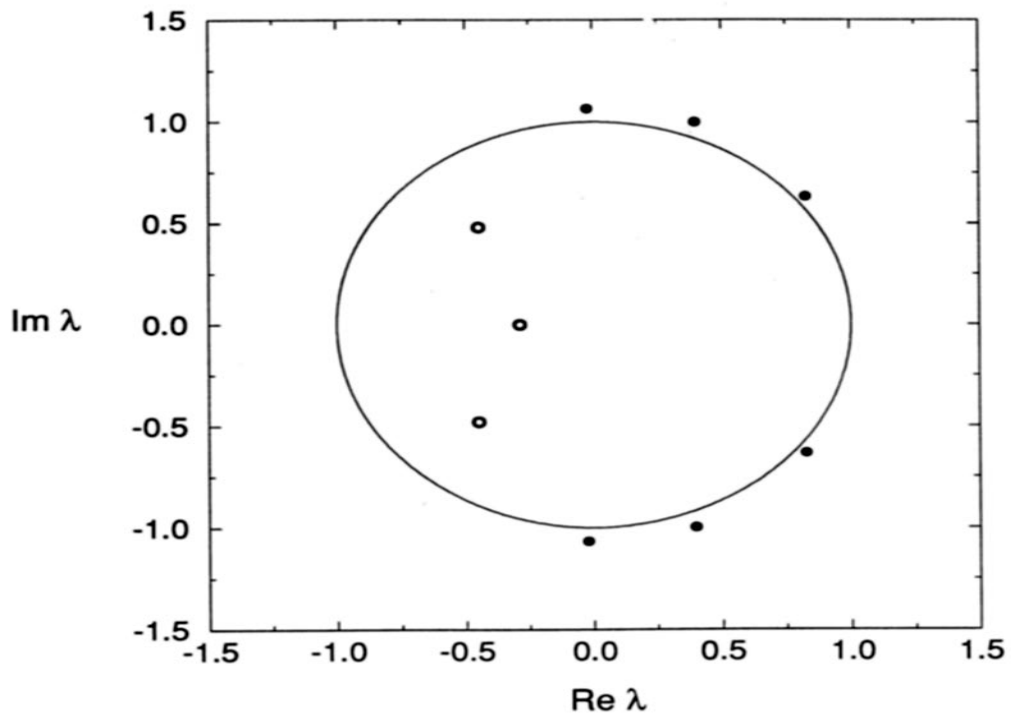
Figure 2.7: Eigenvalues of the unstable four-cell solution calculated as roots of Eq. (2.33). Solid dots represent the unstable eigenvalues of the autonomous system. Open circles show the stable eigenvalues obtained by interrogating the system with random perturbations.

unstable directions. The oscillatory front is monitored by recording the position of the minimum in the front profile. The minimum in the temporal oscillations is then used as the system observable. This choice eliminates the need to construct the n-1 dimensional Poincaré section in time delay coordinates and is therefore convenient for monitoring high-dimensional periodic states of unknown dimensionality. Possible shifts of the attractor[25] do not cause difficulties because such effects are automatically incorporated into Eq. (2.25) from the identification procedure. The bifurcation diagram is shown in Figure 2.8, where the minimum of oscillation is plotted as a function of the reaction zone width. (Further examples of spatiotemporal behavior in the two-cell KS front along with a detailed description of the monitoring technique can be found in Ref. 19).

The two-cell solution exhibits period-1 oscillations at a reaction zone width of $L = 20.7$ where we begin tracking. As the width is decreased, the period-1 orbit becomes unstable through a secondary Hopf bifurcation at $L = 20.57$ with the appearance of quasiperiodic behavior. At $L = 20.4$, the imaginary part of the eigenvalues responsible for the quasiperiodic behavior become zero and two new period-1 solutions appear. As L is decreased further, the quasiperiodic behavior enters the basin of attraction of one of the stable period-1 solutions and nonsymmetric periodic oscillations are exhibited.

The symmetric period-1 solution was stabilized and tracked through the range of L shown in Figure 2.8(a) using the control algorithm with $n = 2$. The controlling perturbation of the boundary condition was found to introduce a negligible displacement along the stable manifolds; therefore, it was necessary to explicitly consider only two unstable degrees of freedom for control. The algorithm was applied according to Eq. (2.29), with small random perturbations added to the control perturbations to interrogate the system. This technique allows the coefficients to be updated every time the bifurcation parameter is changed by repeating the identification procedure. The eigenvalues of the periodic state were calculated from the roots of Eq. (2.12) and are shown in Figure 2.8(b).
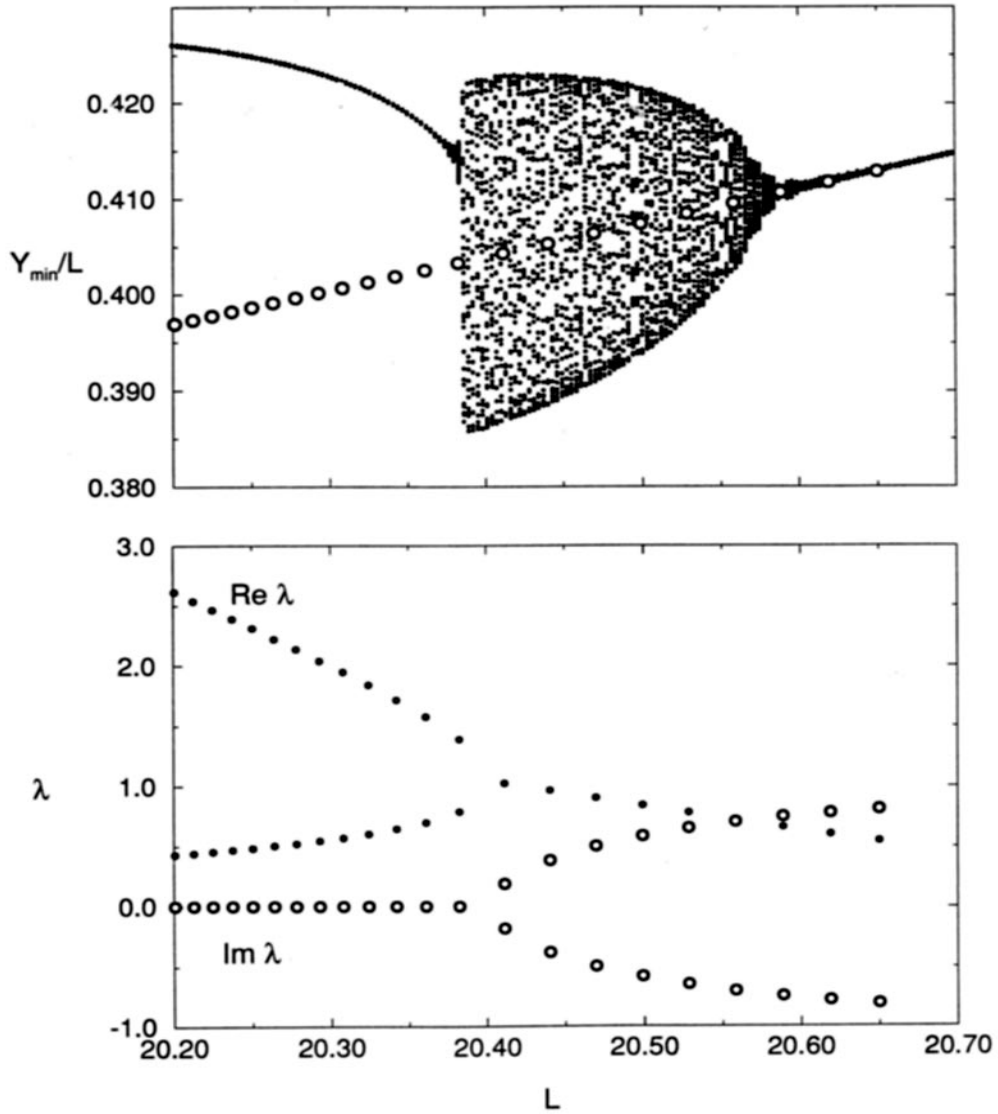
Figure 2.8: Bifurcation diagram showing quasiperiodic behavior of two-cell front of the KS equation. (a) Solid points show minimum of oscillations of front profile minimum. Open circles show symmetrical oscillatory state tracked from right to left through secondary Hopf bifurcation and quasiperiodic oscillations to region where unsymmetrical period-1 oscillation is stable. (b) The real (Re) and imaginary (Im) parts of the eigenvalues as a function of L shown by solid and open circles, respectively.

2.5 Conclusion

A general method for the control of unstable steady or periodic states of dynamical systems has been presented. The algorithm requires only a single observable quantity and acts through perturbations imposed on a single system parameter. For an n dimensional system, n previous observations and $n-1$ previous perturbations are required for control. The states that are stabilized under this control correspond directly to the states of the autonomous system. The algorithm also provides a full characterization of the autonomous state in terms of its effective dimensionality and eigenvalues. The approach can be readily applied to experimental systems without any knowledge of the underlying mechanism or governing equations.

The effective reduction of high dimensional dynamics to a single variable makes the method especially useful for stabilizing spatiotemporal systems by small perturbations localized in space. Local perturbations were sufficient for stabilizing stationary and periodic behavior in the Kuramoto-Sivashinsky equation. We note, however, that this approach may be less successful in spatiotemporal systems with a lower degree of spatial correlation. Local application of the algorithm in such systems will likely result in stabilization only within a correlation radius.

Selection of a particular unstable behavior is also possible using the control algorithm. Setting one of the closed-loop eigenvalues to be the same as the eigenvalue of a chosen unstable manifold will result in a control law that stabilizes all but the selected unstable manifold. This approach requires only very small perturbations and can be used to manipulate the dynamics of a high dimensional system by observing only a single variable. The method may provide a more precise implementation of "anti-control" recently demonstrated in experiments with hippocampal brain tissue.[11]

When coupled with tracking techniques, the algorithm provides a model-independent, path-following method for the bifurcation analysis of experimental systems. The availability of the eigenvalues means that the character of bifurcations in real systems can be determined directly, rather than by inference from observations of qualitative changes in the time series. The algorithm can also be used to extend the parameter range of desired responses, such as stable burning in flame systems or steady output in high dimensional lasers. It should also be noted that even though the

method has been illustrated using a discrete-time approach, it can be reformulated in a continuous-time framework. Such a modification might allow the stabilization of very fast processes by using a control law that is precalculated and then implemented with an analog circuit.

2.6 References

[1] T. Shinbrot, C. Grebogi, E. Ott, and J. A. Yorke, "Using Small Perturbations to Control Chaos," *Nature* **363**, 411-417 (1993).

[2] W. L. Ditto, S. N. Rauseo, and M. L. Spano, "Experimental Control of Chaos," *Phys. Rev. Lett.* **65**, 3211-3214 (1990).

[3] E. R. Hunt, "Stabilizing High-Period Orbits in a Chaotic System: The Diode Resonator," *Phys. Rev. Lett.* **67**, 1953-1955 (1991).

[4] T. Carroll, I. Triandaf, I. B. Schwartz, and L. Pecora, "Tracking Unstable Orbits Experimentally," *Phys. Rev. A* **46**, 6189-6192 (1992).

[5] R. Roy, T. W. Murphy, T. D. Maier, Z. Gills, and E. R. Hunt, "Dynamical Control of a Chaotic Laser: Experimental Stabilization of a Globally Coupled System," *Phys. Rev. Lett.* **68**, 1259-1262 (1992).

[6] S. Bielawski, D. Derozeir, and P. Glorieux, "Experimental Characterization of Unstable Periodic Orbits by Controlling Chaos," *Phys. Rev. A* **47**, R2492-R2495 (1993).

[7] V. Petrov, V. Gáspár, J. Masere, and K. Showalter, "Controlling Chaos in the Belousov-Zhabotinsky Reaction," *Nature* **361**, 240-243 (1993).

[8] P. Parmananda, P. Sherard, R. W. Rollins, and H. D. Dewald, "Control of Chaos in an Electrochemical Cell," *Phys. Rev. E* **47**, R3003-R3006 (1993).

[9] F. W. Schneider, R. Blittersdorf, A. Förster, T. Hauck, D. Lebender, and J. Müller, "Control of Dynamic States with Time Delay Between 2 Mutually Flow Rate Coupled Reactors," *J. Phys. Chem.* **97**, 12244-12248 (1993).

[10] A. Garfinkel, M. L. Spano, W. L. Ditto, and J. N. Weiss, "Controlling Cardiac Chaos," *Science* **257**, 1230-1235 (1992).

[11] S. J. Schiff, K. Jerger, D. H. Duong, T. Chang, M. L. Spano, and W. L. Ditto, "Controlling Chaos in the Brain," *Nature* **370**, 615-620 (1994).

[12] E. Ott, C. Grebogi, and J. A. Yorke, "Controlling Chaos," *Phys. Rev. Lett.* **64**, 1196-1199 (1990).

[13] K. Pyragas, "Continuous Control of Chaos by Self-Controlling Feedback," *Phys. Lett. A* **170**, 421-428 (1992).

[14] K. Pyragas and A. Tamasevicius, "Experimental Control of Chaos by Delayed Self-Controlling Feedback," *Phys. Lett. A* **180**, 99-102 (1993).

[15] P. M. Alsing, A. Gavrielides, and V. Kovanis, "History Dependent Control of Unstable Periodic Orbits," *Phys. Rev. E* **50**, 1968-1977 (1994).

[16] B. Peng, V. Petrov, and K. Showalter, "Controlling Chemical Chaos," *J. Phys. Chem.* **95**, 4957-4959 (1991).

[17] V. Petrov, B. Peng, and K. Showalter, "A Map-Based Algorithm for Controlling Low-Dimensional Chaos," *J. Chem. Phys.* **96**, 7506-7513 (1992).

[18] V. Petrov, M. F. Crowley, and K. Showalter, "Tracking Unstable Periodic Orbits in the Belousov-Zhabotinsky Reaction," *Phys. Rev. Lett.* **72**, 2955-2958 (1994).

[19] V. Petrov, M. F. Crowley, and K. Showalter, "Controlling Spatiotemporal Dynamics of Flame Fronts," *J. Chem. Phys.* **101**, 6606-6615 (1994).

[20] S. Bielawski, D. Derozeir, and P. Glorieux, "Comtrolling Unstable Periodic Orbits by a delayed continuous feedback," *Phys. Rev. E* **49**, R971-R794 (1994).

[21] Z. Gills, C. Iwata, R. Roy, I. B. Schwartz, and I. Triandaf, "Tracking Unstable Steady States: Extending the Stability Regime of a Multimode Laser System," *Phys. Rev. Lett.* **69**, 3169-3172 (1992).

[22] A. Hjelmfelt and J. Ross, *J. Phys. Chem.* "Experimental Stabilization of Unstable Steady States in Oscillatory and Excitable Reaction Systems," **94**, 1176-1179 (1994).

[23] S. Bielawski, M. Bouazaoui, D. Derozeir, and P. Glorieux, "Stabilization and Characterization of Unstable Steady States in a Laser," *Phys. Rev. A* **47**, 3276-3279 (1993).

[24] R. W. Rollins, P. Parmananda, and P. Sherard, "Controlling Chaos in Highly Dissipative Systems: a Simple Recursive Algorithm," *Phys. Rev. E* **47**, R780-R783 (1993).

[25] U. Dressler and G. Nitsche, "Controlling Chaos Using Time Delay Coordinates," *Phys. Rev. Lett.* **68**, 1964 (1992).

[26] D. Auerbach, "Controlling Extended Systems of Chaotic Elements," *Phys. Rev. Lett.* **72**, 1184-1187 (1994).

[27] Hu Gang, and Qu Zhilin, "Controlling Spatiotemporal Chaos in Coupled Map Lattice Systems," *Phys. Rev. Lett.* **72**, 68-71 (1994).

[28] D. Auerbach, C. Grebogi, E. Ott, and J. A. Yorke, "Controlling Chaos in High Dimensional Systems," *Phys. Rev. Lett.* **69**, 3479-3482 (1992).

[29] F. J. Romeiras, C. Grebogi, E. Ott, and W.P. Dayawansa, "Controlling Chaotic Dynamical Systems," *Physica D* **58**, 165-192 (1992).

[30] G. C. Goodwin and K. S. Sin, *Adaptive Filtering, Prediction, and Control* (Prentice-Hall, Englewood Cliffs, N.J. 1984).

[31] L. Ljung, *System Identification - Theory for the User*, (Prentice-Hall, Englewood Cliffs, N.J. 1987).

[32] J.C. Dole, K. Glover, P. Khargonekar, and B. Francis, *IEEE Trans. Automat. Contr.* **34**, 1831 (1989).

[33] G.I. Sivashinsky, "Instabilities, Pattern Formation and Turbulence in Flames," *Ann. Rev. Fluid Mech.* **15**, 179 (1983).

Chapter 3

Normal Modes for Chemical Reactions from Time Series Analysis

3.1 Introduction

The dynamical behavior of a chemical reaction close to a stationary state is described by the elements of the associated Jacobian matrix. If a fully specified model is known for a system, it is straightforward to calculate the stationary state and the corresponding Jacobian matrix. However, the reverse problem of deducing the mechanism from a knowledge of the Jacobian is much more difficult, and, in addition, such problems do not afford unique solutions. Nevertheless, valuable information about the mechanism can be gleaned from a knowledge of the Jacobian matrix.[1] In fact, mechanistic insights can be obtained even if only a part of the matrix or just the signs of the elements are known.[2-6] Most of the methods that have been proposed for experimental determination of the Jacobian require the ability to measure the concentrations of all species that significantly affect the dynamics. Unfortunately, this condition is often not met in practice.

A previously developed method, called quenching analysis,[7] overcomes many of these difficulties by carrying out the chemical reaction in a continuous-flow stirred tank reactor (CSTR) close to a supercritical Hopf bifurcation. The operating conditions are adjusted such that the system shows small amplitude, sinusoidal limit

cycle oscillations around a saddle focus. For each of the essential species there exists a characteristic perturbation magnitude and phase of the oscillation for which the oscillatory behavior can be temporarily stopped (quenched) by the effectively instantaneous addition of an appropriate compound. The perturbation moves the current state of the system from the limit cycle to the codimension-two stable manifold of the saddle focus, from which it slowly returns to the limit cycle. This behavior is universal for a chemical system near a supercritical Hopf bifurcation. Observing the concentration of just one of the essential species and measuring the quenching parameters for n essential species allows the calculation of the $2n$ elements of the two left eigenvectors associated with the Hopf oscillatory mode. The corresponding right eigenvectors of the Jacobian matrix can be determined by additional measurements of the oscillating concentrations of $n-3$ of the remaining species.[8] The concentrations of the stationary state can then be computed from an additional quenching experiment performed by an instantaneous dilution of the system. The quenching method has been used to investigate the mechanisms of a number of different chemical systems.[9]

In this chapter we present a new method for determining the complete Jacobian matrix of an n-species chemical system in a stable stationary state. The method consists of making n time series measurements of a single property that is a linear (but not necessarily known) function of the concentrations. In each series of measurements, a different set of species is subjected to random known perturbations at regular time intervals. If some of the characteristic times are much smaller than the time interval of the measurements, it is then not possible to determine the complete Jacobian; however, the method allows the computation of the characteristic times of the remaining slow modes together with the corresponding left eigenvectors of the Jacobian matrix.

3.2   Method

Following linear control theory,[10] we write an autoregression equation for successive measurements and perturbations of a chemical system very close to a stationary state.[11] The measured quantity is typically some linear function of the concentrations of the dynamical species, such as optical absorbance. Considering an n-dimensional

chemical system that is perturbed at equal time intervals by the addition of one (or more) of the dynamically important species, the autoregression equation takes the form:

$$
\begin{aligned}
y_k &= a_1 y_{k-1} + a_2 y_{k-2} + \ldots + a_n y_{k-n} + b_0 + \\
&\quad b_1 w_{k-1} + b_2 w_{k-2} + \ldots + b_n w_{k-n}
\end{aligned} \tag{3.1}
$$

where y_k is the measured quantity and w_k is the magnitude of the perturbation at time t_k. This equation is often seen without the bias term, b_0, which can always be set to zero by a suitable scaling for the y measurements. We find it convenient, however, to consider b_0 as a parameter that is identified from the experimental data in the same step as a_i and b_i. It is necessary to carry out at least $3n + 1$ readings of y in order to obtain the $2n + 1$ equations needed to determine the $2n + 1$ coefficients. However, in practice, many more readings are typically required to adequately determine the coefficients, due to unavoidable noise. The method of singular value decomposition (SVD)[12, 13] allows the coefficients to be accurately estimated from the over-determined system. The magnitudes of the perturbations, w_k, are randomly varied in order to ensure that the equations determining the coefficients a_i and b_i can be solved. Alternatively, perturbations with the same magnitude can be made at irregular intervals.

In Figure 3.1, we show a specific example of the identification procedure. We use time series generated by integrating the Oregonator model[14] of the Belousov-Zhabotinsky[15] reaction (BZ), with parameters adjusted to fit results from previous quenching experiments.[16, 17] The model parameters were chosen so that the chemical system exhibits a stable stationary state. The Oregonator model includes three essential chemical species, $HBrO_2$, Br^-, and Ce^{4+}, designated X, Y, and Z, with concentrations X, Y, and Z. The eigenvalues of the Jacobian at the stationary point are -0.34 and $-0.012 \pm i\,0.087$. The upper curve shows the evolution of the concentration Z in response to the perturbations shown in the lower part of the figure. The circles indicate the sampling intervals. The coefficients of the autoregression eq (3.1) for the three-dimensional model were determined from 500 readings of Z and the corresponding random perturbations at 3.0 s intervals. The singular value decomposition yields the following values:[18]
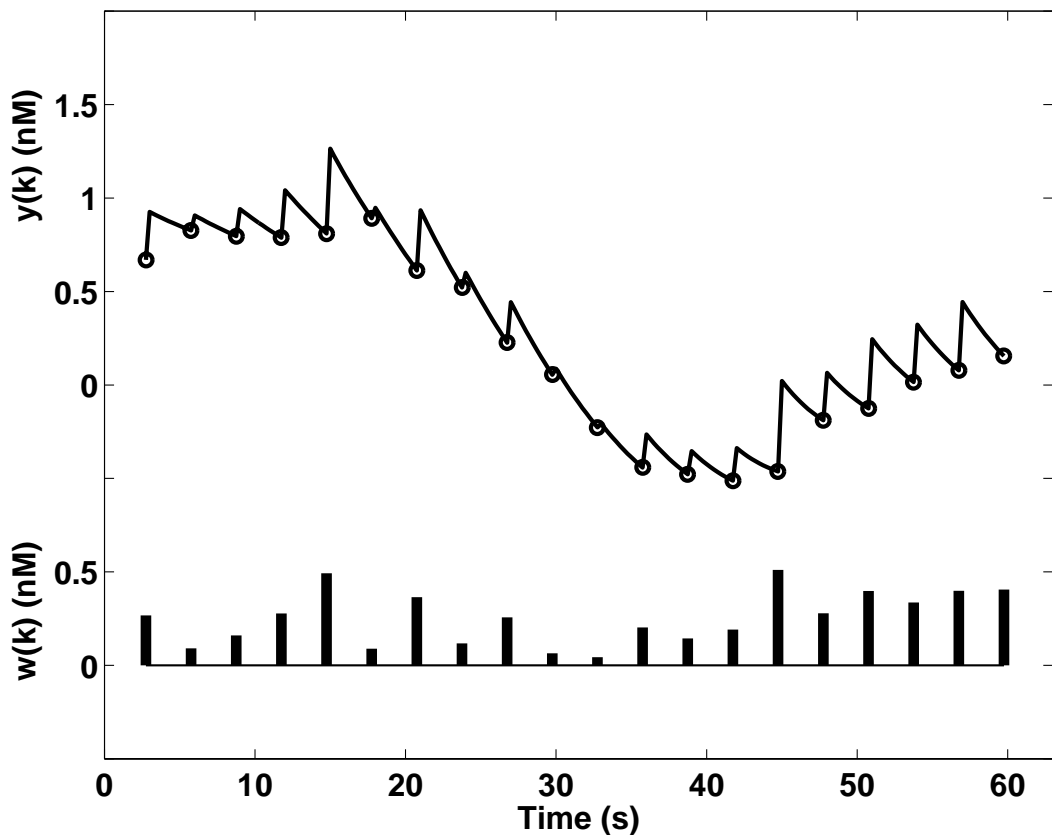
Figure 3.1: Perturbations w_k and time series of responding chemical system $y_k = (Z - Z_s)$. Calculations were carried out using the three-variable Oregonator model, identical to model M3 described in Nielsen *et al.*[16] The system is perturbed by making momentary changes in the variable $Z = [\mathrm{Ce^{4+}}]$. The perturbations are random in magnitude, drawn from a uniform distribution in the interval from 0.0 to 0.5 nM. The measurements of Z are made at 3.0 s intervals (indicated by circles), and a perturbation is made immediately after each measurement. All parameters are the same as in Nielsen *et al.*[16] except $[\mathrm{H^+}]$ is changed from 1.0 M to 0.8 M to produce a weakly stable system. The stationary state concentrations for the three variable species are $X_s = 2.84 \times 10^{-8}$ M, $Y_s = 2.17 \times 10^{-7}$ M, and $Z_s = 1.37 \times 10^{-7}$ M.

$$(a_1, a_2, a_3) = (2.215, -1.586, 0.338)$$

$$(b_1, b_2, b_3) = (0.589, -1.031, 0.328)$$

In simulations, the adequacy of the coefficients obtained from the autoregression model can be readily checked by comparing the predicted signal with the original computed signal. The residual error is quite small, with the one-step prediction of Z by the autoregression model virtually indistinguishable from the time series shown in Figure 3.1.

The numerical experiment was carried out three times, with perturbations applied, in turn, to the X, Y, and Z concentration variables of the model. The measured variable in each case was Z. The range of perturbation magnitudes for each run was chosen to yield responses of similar magnitude in Z. For the perturbations in X, Y, and Z, the maximum magnitudes were 1.7 x 10^{-10} M, 3.9 x 10^{-10} M, and 5.2 x 10^{-10} M, respectively. The coefficients of the associated autoregression models were then used to determine the elements of the Jacobian matrix as described below.

3.3   Computing the Jacobian Matrix

In a series of measurements and perturbations, the perturbation at time t_k results in a shift in concentration space of $w_k\,\mathbf{g}$ in a fixed direction given by the vector $\mathbf{g}$. The coefficients a_j and b_j of the autoregression equation (3.1) for the series can now be used to write the following matrices:

$$\mathbf{L} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ a_3 & a_2 & a_1 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} 1 & 0 & 0 \\ -a_1 & 1 & 0 \\ -a_2 & -a_1 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

The coefficients a_j are characteristic of the dynamics of the unperturbed system and are independent of the direction of perturbation, whereas the coefficients b_j depend on $\mathbf{g}$. By carrying out three different series with perturbations in three directions, given by the linearly independent vectors, $\mathbf{g}_1$, $\mathbf{g}_2$, and $\mathbf{g}_3$, we obtain three different vectors, $\mathbf{b}_1$, $\mathbf{b}_2$, and $\mathbf{b}_3$. These vectors can be combined in matrices $\mathbf{G} = (\mathbf{g_1}, \mathbf{g_2}, \mathbf{g_3})$

and $\mathbf{B} = (\mathbf{b_1}, \mathbf{b_2}, \mathbf{b_3})$, with the vectors $\mathbf{g}$ and $\mathbf{b}$ as columns. Using $\mathbf{G}$ and $\mathbf{B}$, we can now compute the matrix $\mathbf{F}$ governing the discrete dynamics corresponding to a fixed sampling interval Δt,

$$\mathbf{F} = \mathbf{G}\,\mathbf{B}^{-1}\,\mathbf{C}\,\mathbf{L}\,\mathbf{C}^{-1}\,\mathbf{B}\,\mathbf{G}^{-1} \tag{3.2}$$

and hence the Jacobian matrix from the relation

$$\mathbf{F} = e^{\mathbf{J}\Delta t} \tag{3.3}$$

The development of eqs (3.2) and (3.3) can be found in the Appendix.

3.4   Results

Repeating the integration of the Oregonator model, first with Br^- (Y) and then with $HBrO_2$ (X) replacing Ce^{4+} (Z) as the perturbed species, we obtain the same a coefficients but different b coefficients. This procedure makes $\mathbf{G}$ equal to the 3D identity matrix, since in each case the perturbation vector, $\mathbf{g}$, is directed along one of the coordinate axes in concentration space.

Forming the matrices $\mathbf{L}, \mathbf{C}, \mathbf{B}$, and $\mathbf{G}$ and inserting them into eqs (3.2) and (3.3) results in the following Jacobian matrix (in units of s^{-1}):

$$\mathbf{J} = \begin{pmatrix} -0.121 & -0.052 & 0.000 \\ -0.530 & -0.084 & 0.131 \\ 0.811 & 0.000 & -0.165 \end{pmatrix}$$

in good agreement with the Jacobian matrix calculated directly from the Oregonator model:

$$\mathbf{J} = \begin{pmatrix} -0.1174 & -0.0528 & 0.0000 \\ -0.5202 & -0.0835 & 0.1320 \\ 0.8064 & 0.0000 & -0.1670 \end{pmatrix}$$

In an actual experiment, the signal to noise ratio (SNR) is limited due to uncontrollable external fluctuations. It is therefore important to assess the robustness of

the method in the presence of noise. Figure 3.2 shows the relative Jacobian error as a function of the SNR. We see that the error increases proportionally with the noise level above a certain level. It is important to note that the effect of noise on the reconstructed Jacobian error depends sensitively on the eigenvalues and a degradation of performance occurs even at low noise levels in very stiff systems. A small amount of noise (0.1 % of the signal amplitude) was added to the measured variable in each of the following calculations in order to simulate realistic experimental measurements.

The Jacobian error reflects noise in the data and nonlinearities of the system, which can, to some extent, be separated by varying the amplitude of the perturbations. Excessively large perturbations result in deviations from the linear regime and degrade the precision of the reconstructed Jacobian. Perturbations that are too small, on the other hand, reduce the available signal to noise ratio and are also undesirable. These trends can be seen in Figure 3.3, which shows the relative Jacobian error as a function of perturbation amplitude. We see that the perturbation amplitude can be as large as 5.2 x 10^{-9} M, which is about 4.0 % of the stationary state concentration (Z_s = 1.37 x 10^{-7} M), before the relative Jacobian error begins to increase. Perturbation amplitudes of about 1.0 % of the stationary state concentration give rise to a relative Jacobian error of about 3.0 %.

There is similarly an optimal choice of the sampling interval, as shown in Figure 3.4. The faster dynamical modes become unobservable for very long sampling intervals, while very short sampling intervals give rise to measured signals primarily reflecting the noise. The optimal value of the sampling interval is therefore a complex function of the system time scales and properties of the noise. The variation in the coefficients during a particular experiment or from experiment to experiment can serve as an empirical measure of the robustness of the autoregression model. We see in Figure 3.4 that a 3.0 s sampling interval is close to the optimal value, which is comparable to the reciprocal eigenvalues of $\mathbf{J}$. It should also be noted that a large variation in the Jacobian error was observed between individual runs. It is frequently possible to obtain errors as low as 1.0 % with an optimal choice of parameters, but large errors occasionally occur, which limits the average precision to approximately 3.0 %.

Reconstruction of the Jacobian depends upon the precision of the autoregression model obtained from the experimental data. Increasing the number of samples results
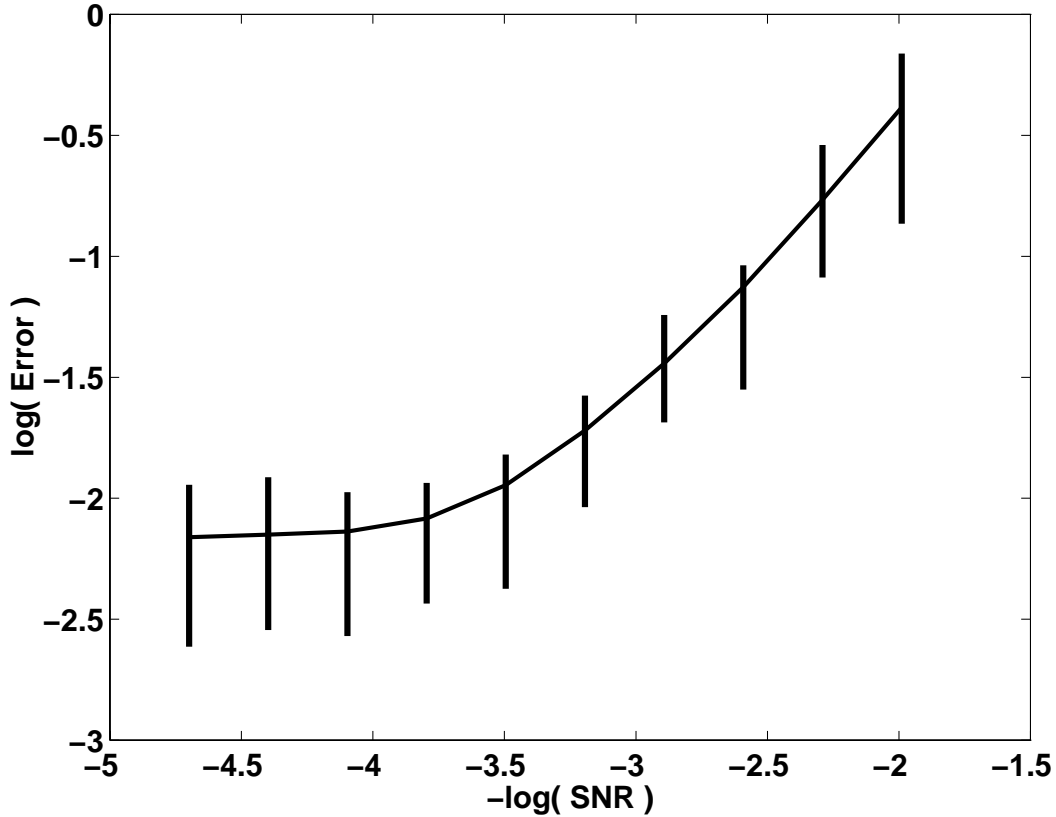
Figure 3.2: Dependence of the relative Jacobian error on the signal to noise ratio. Uniform noise is added to the measurements of y_k and the relative Jacobian error is shown as a function of relative noise amplitude. The vertical lines show the standard deviation from the mean values for 40 numerical experiments. The abscissa is the logarithm of the rms of the noise relative to the rms of the signal. The relative Jacobian error is defined as the Euclidian norm of the difference between the analytical and reconstructed Jacobians divided by the Euclidian norm of the analytical Jacobian, where the Euclidian norm is the square root of the sum of the squared elements. Parameters and procedures are the same as in Fig. 3.1.

Figure 3.3: Dependence of the relative Jacobian error on the perturbation amplitude. The solid curve shows the error as a function of the maximum perturbation $\|w\|$ relative to $\|w_0\| = 0.5$ nM as in Fig. 1. The perturbation amplitudes for X and Y were varied proportionally, as each Jacobian reconstruction requires data from three experiments, with each interrogating one of the variables. The vertical lines show the standard deviation from the mean values for 40 numerical experiments. Low level noise (0.1 % of the signal amplitude) was added to the measured variable. The parameters and procedures are the same as those for Fig. 3.1. The relative Jacobian error is defined in Fig. 3.2.
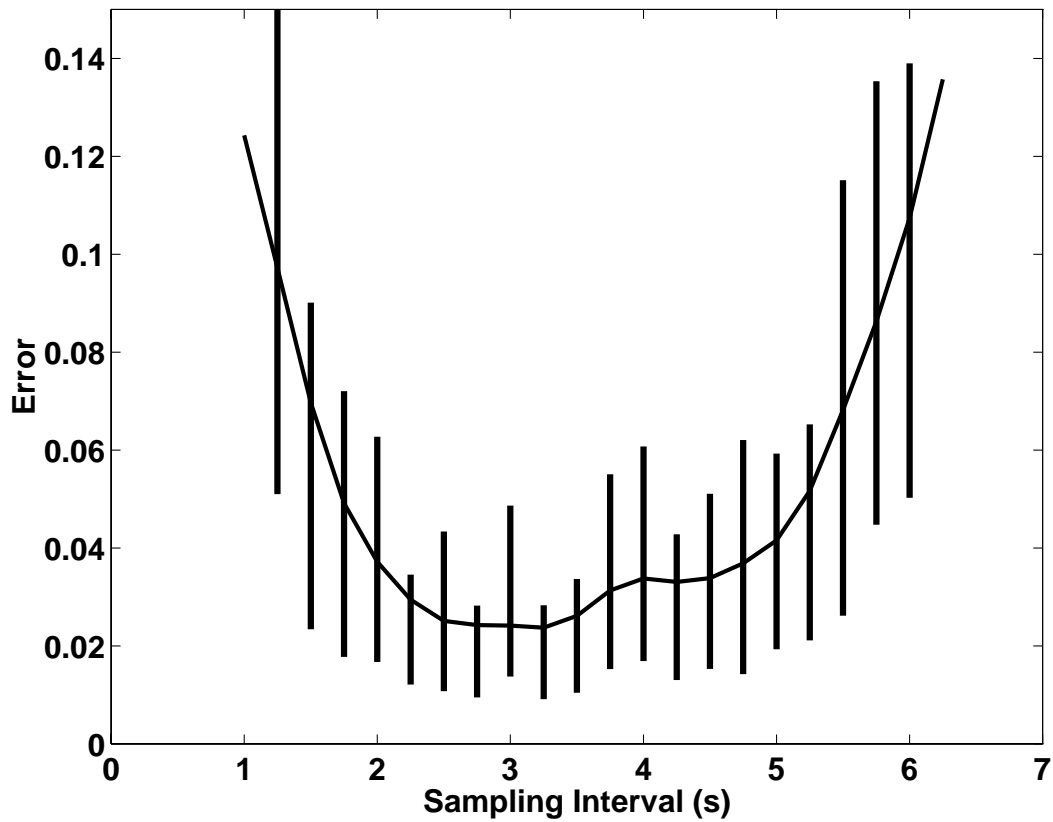
Figure 3.4: Dependence of the relative Jacobian error on the sampling interval. The solid curve shows the error as a function of the sampling interval, and the vertical lines show the standard deviation from the mean values for 40 numerical experiments. Low level noise (0.1% of the signal amplitude) was added to the measured variable. Parameters and procedures are the same as in Fig. 3.1, except that the perturbation amplitude was adjusted to maintain a constant amplitude of the response signal. The relative Jacobian error is defined in Fig. 3.2.
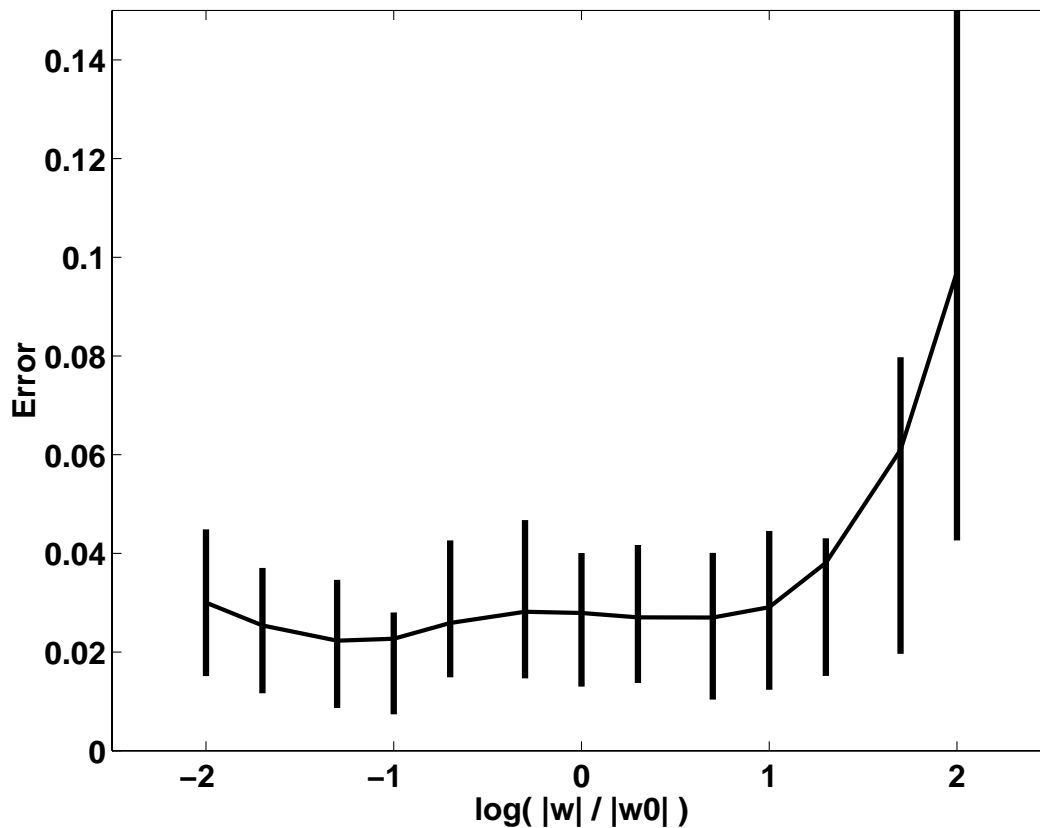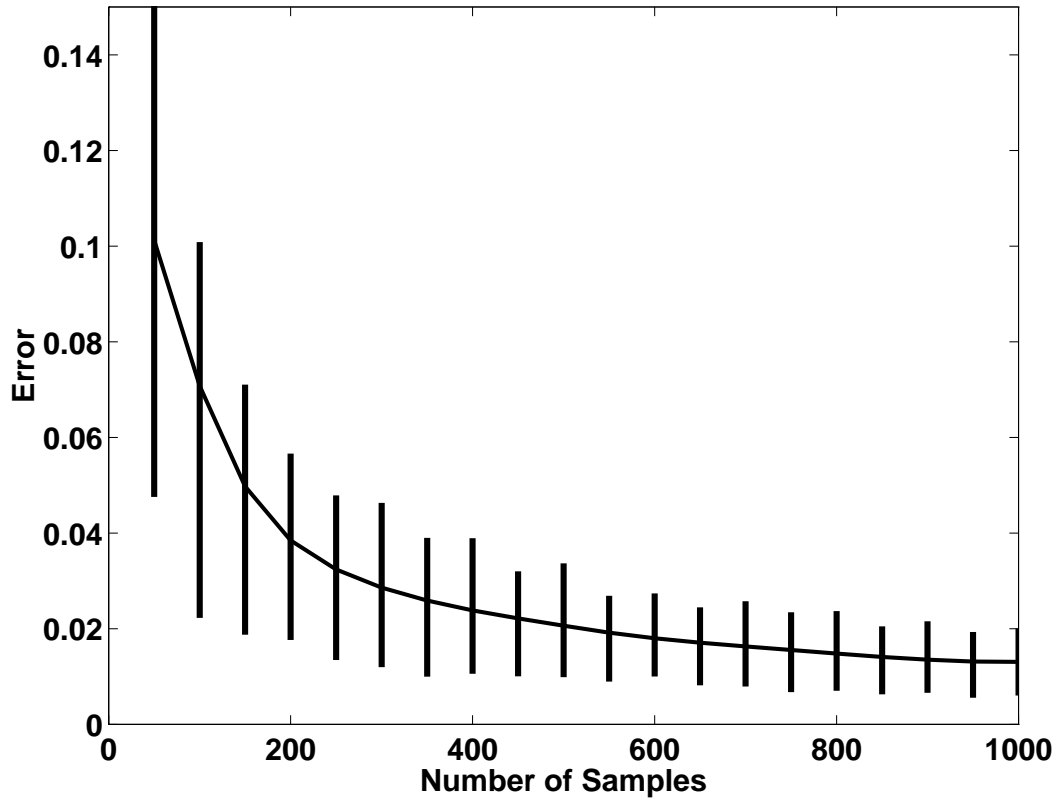
Figure 3.5: Dependence of the relative Jacobian error on the number of samples. The solid curve shows the error as a function of the total number of samples used with each dynamical variable in reconstructing the Jacobian. The vertical lines show the standard deviation from the mean values for 40 numerical experiments. Low level noise (0.1 % of the signal amplitude) was added to the measured variable. Parameters and procedures are the same as in Fig. 3.1. The relative Jacobian error is defined in Fig. 3.2.

in a more precise autoregression model. Using the same parameters as above, we show in Figure 3.5 the relative Jacobian error as a function of the number of data pairs collected in each numerical experiment (for each dynamical variable). We see in this particular example that there is no significant advantage in increasing the number of samples above approximately 500 data pairs. We have not determined how this error dependence varies with different dynamical systems.

We note that the identification procedure can be carried out in an alternative manner that may be advantageous in actual experiments. Rather than random amplitude perturbations delivered at constant time intervals, constant amplitude perturbations can be delivered at nonperiodic time intervals. Identical values for the coefficients (a_1, a_2, a_3) and (b_1, b_2, b_3) were obtained when the calculation for Figure 3.1 was repeated using perturbations at time intervals that were $\frac{t_k}{3.0s}$ multiples of 7, 13, or 17. This scheme has the advantage that constant perturbations can be delivered more accurately and conveniently in an experimental setting. In addition, fewer perturbations are required for the system identification, and the system state is therefore moved away from the stationary state less than it would be with constant time interval perturbations.

3.5   Discussion

The determination of chemical reaction mechanisms is often difficult, particularly for complex reactions, because no systematic methods exist for identifying the essential component steps of a reaction. Chemical mechanisms are usually deduced by considering combinations of elementary steps that give rise to a scheme consistent with the experimentally measured chemical kinetics of the reaction. Relaxation kinetics offers a convenient approach for experimentally probing chemical mechanisms. Originally developed by Eigen and co-workers,[19] the technique involves monitoring relaxations very close to the equilibrium state such that the reaction follows first-order kinetics. Relaxation methods have been utilized extensively for the investigation of chemical mechanisms, particularly for very fast reactions.[20] Recently, there has been renewed interest in perturbation methods for mechanism elucidation, in which the dynamical behavior around a nonequilibrium stationary state is probed. Quenching techniques,

where an oscillatory reaction is driven to its unstable stationary state by precisely time perturbations, have been successfully applied to several experimental chemical systems.[7–9] These studies have focused on the Jacobian matrix of the nonequilibrium stationary state to provide insights into the possible restrictions of a scheme of mechanistic steps. The elements of the Jacobian matrix tell how each dynamical species responds to all of the other dynamical species, as well as to itself, and therefore provides vital information on allowed and forbidden steps in a chemical mechanism.[1]

Determining the Jacobian matrix from experimental data and deducing a reaction mechanism that is consistent with the Jacobian elements is an important challenge. A number of studies have recently considered this problem in efforts to develop a systematic approach for mechanism elucidation.[2–6] In this chapter, we have addressed the essential step of determining the Jacobian matrix from experimental data. Our approach is based on classical linear control theory, where a system in a stationary or periodic state is subjected to very small, random perturbations to determine the "normal modes" of the reaction. The perturbations and corresponding responses form a discrete time series as a sequence of data pairs that can be analyzed to yield the Jacobian matrix.

In principle, the Jacobian can be unambiguously determined from time series analysis of a chemical reaction subjected to random perturbations, provided that the system is observable and controllable.[21] In practice, however, there are uncertainties that pose technical challenges to the method. The first is ascertaining all of the dynamically important species and then devising a means to impose perturbations on each of these species. As we have noted, the method yields useful mechanistic information even if some of the essential species cannot be accessed. The other limitation is the effect of noise on the accuracy of the Jacobian matrix elements when applying the method in an experimental setting. For some systems, the method is quite robust to external noise; however, for other systems, especially those with a large negative eigenvalue, noise may significantly degrade the reliability of the Jacobian matrix. It should still be possible in such cases to determine the sign of the Jacobian matrix elements, which can offer valuable insights into the mechanistic possibilities of a chemical reaction. There are other well-known controllability and observability pathologies, such as degenerate eigenvalues, that cause the method to fail.[21] On

the other hand, in favorable cases the method can offer an experimental means to determine the possible chemical steps as well as the associated rate constants.

We have shown that the Jacobian matrix of a chemical system in a stable stationary state can be determined from the time series of a single observable variable by applying perturbations to the relevant chemical species. If some of the modes of the system are too fast to be measured by the method, the information on the remaining modes can still be recovered. The performance of the method depends upon an appropriate choice of the perturbation magnitude and sampling interval, and, in general, increases with an increasing amount of data available for processing. The method can be readily applied to higher-dimensional chemical systems by a straightforward extension of the corresponding vectors and matrices.

3.6   Appendix

The kinetic equations for a homogeneous chemical system with n chemical species can be written as

$$\frac{d\mathbf{c}(t)}{dt} = \mathbf{f}(\mathbf{c}(t)) \tag{3.4}$$

where the column vector $\mathbf{c}(t)$ describes the time dependent concentrations. We assume that the system has a stationary state, $\mathbf{c_s}$. A linear expansion of the kinetic equations around the stationary state gives

$$\frac{d\mathbf{u}(t)}{dt} = \mathbf{J} \cdot \mathbf{u}(t) + \dots \tag{3.5}$$

where $\mathbf{u}(t) = \mathbf{c}(t) - \mathbf{c_s}$ describes the deviation from the stationary state and $\mathbf{J}$ is the Jacobian matrix with elements $J_{ij} = \left.\frac{\partial f_i(\mathbf{c})}{\partial c_j}\right|_{\mathbf{c}=\mathbf{c_s}}$. We consider values of $\mathbf{u}(t)$ at a set of equidistant discrete moments t_k. The vectors $\mathbf{u}(t_k)$ and $\mathbf{u}(t_{k-1})$ are formally related by the expression

$$\mathbf{u}(t_k) = e^{\mathbf{J}\Delta t}\mathbf{u}(t_{k-1}), \tag{3.6}$$

where $\Delta t = t_k - t_{k-1}$. The matrix $\mathbf{F} = e^{\mathbf{J}\Delta t}$ has the same eigenvectors as $\mathbf{J}$ and has eigenvalues $\rho_j = e^{\lambda_j \Delta t}$, where λ_j are the eigenvalues of $\mathbf{J}$.

The evolution of the chemical system is monitored by measuring a property $y(t)$, which is a function of $\mathbf{c}(t)$. A linear expansion of y from $y_s = y(\mathbf{c_s})$ gives

$$z = y - y_s = \mathbf{h^T} \cdot \mathbf{u}, \tag{3.7}$$

where $\mathbf{h}$ is the vector of coefficients of the first-order terms in the expansion. All essential changes in the system are reflected in $y(t)$, provided that the matrix

$$\mathbf{A} = \begin{pmatrix} \mathbf{h^T} \\ \mathbf{h^T} \cdot \mathbf{F} \\ \mathbf{h^T} \cdot \mathbf{F^2} \\ \vdots \\ \mathbf{h^T} \cdot \mathbf{F^{n-1}} \end{pmatrix} \tag{3.8}$$

is of rank n (the observability condition).[21]

We follow Lee[21] in the subsequent transformations leading to the identification of the Jacobian matrix elements. Perturbations are applied to the system at t_k resulting in an instantaneous concentration change of $\mathbf{g}w_k$. To ensure that the perturbation excites all of the characteristic modes, the vector $\mathbf{g}$ must be selected such that the matrix

$$(\mathbf{g}, \mathbf{F} \cdot \mathbf{g}, \mathbf{F^2} \cdot \mathbf{g}, \ldots, \mathbf{F^{n-1}} \cdot \mathbf{g}) \tag{3.9}$$

has rank n (the controllability condition).[21] By convention, the perturbation at any time t_k is made immediately after the measurement and therefore does not affect the value measured at t_k. Successive $\mathbf{u}_k$ and $z_k = z(t_k)$ are related by

$$\begin{aligned} \mathbf{u}_k &= \mathbf{F} \cdot (\mathbf{u}_{k-1} + \mathbf{g}w_{k-1}) \\ &= \mathbf{F} \cdot \mathbf{u}_{k-1} + \mathbf{F} \cdot \mathbf{g}w_{k-1} \\ z_k &= \mathbf{h^T} \cdot \mathbf{u_k}. \end{aligned} \tag{3.10}$$

By direct substitution it can be shown that the coordinate transformation $\mathbf{v} = \mathbf{A} \cdot \mathbf{u}$ transforms eq (3.10) into the canonical form

$$\mathbf{v}_k = \mathbf{L} \cdot \mathbf{v}_{k-1} + \mathbf{d}w_{k-1} \tag{3.11}$$

$$z_k = (1, 0, \cdots, 0) \cdot \mathbf{v}_k$$

where

$$\mathbf{L} = \mathbf{A} \cdot \mathbf{F} \cdot \mathbf{A}^{-1} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ L_{n1} & L_{n2} & L_{n3} & \cdots & L_{nn} \end{pmatrix} \tag{3.12}$$

$$\mathbf{d} = \mathbf{A} \cdot \mathbf{F} \cdot \mathbf{g}$$

The elements L_{ni} of the matrix $\mathbf{L}$ can be determined explicitly from $\mathbf{F}$ and $\mathbf{h}$, but L_{ni} can also be determined from the following autoregression equation involving the experimental observables y_k and perturbations w_k:

$$y_k = a_1 y_{k-1} + a_2 y_{k-2} + \ldots + a_n y_{k-n} + b_0 +$$
$$b_1 w_{k-1} + b_2 w_{k-2} + \ldots + b_n w_{k-n}. \tag{3.13}$$

It can be shown[21] that

$$L_{n1} = a_n, L_{n2} = a_{n-1}, \ldots, L_{nn} = a_1 \tag{3.14}$$

and that the elements of $\mathbf{d}$ are related to the coefficients of eq (3.13) by

$$\mathbf{d} = \mathbf{C}^{-1} \cdot \mathbf{b} \tag{3.15}$$

where

$$\mathbf{C} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -a_1 & 1 & 0 & \cdots & 0 \\ -a_2 & -a_1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ -a_{n-1} & -a_{n-2} & -a_{n-3} & \cdots & 1 \end{pmatrix} \text{ and } \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{pmatrix} \tag{3.16}$$

In an experimental investigation, $\mathbf{F}$ and, consequently, $\mathbf{A}$ are usually unknown. The coefficients a_i and b_i in eq (3.13) can be identified from time series. The matrices $\mathbf{L}$ and $\mathbf{C}$ are determined by these coefficients, and $\mathbf{h}$ and $\mathbf{g}$ are defined by the experimental conditions. We see from eq (3.12) and eq (3.15) that $\mathbf{A}$ must satisfy the equation $\mathbf{A} \cdot \mathbf{F} \cdot \mathbf{g} = \mathbf{C}^{-1} \cdot \mathbf{b}$. By carrying out perturbation experiments with n linearly independent vectors $\mathbf{g}$ and determining the corresponding vectors $\mathbf{b}$, the following matrix equation can be established:

$$\mathbf{A} \cdot \mathbf{F} \cdot \mathbf{G} = \mathbf{C}^{-1} \cdot \mathbf{B} \tag{3.17}$$

where $\mathbf{G} = (\mathbf{g}_1, \ldots, \mathbf{g}_n)$ and $\mathbf{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$. Introducing

$$\boldsymbol{\Phi} = \mathbf{A} \cdot \mathbf{F} = \mathbf{C}^{-1} \cdot \mathbf{B} \cdot \mathbf{G}^{-1} \tag{3.18}$$

the $\mathbf{F}$ matrix can be determined from

$$\mathbf{F} = \boldsymbol{\Phi}^{-1} \cdot \mathbf{L} \cdot \boldsymbol{\Phi} \tag{3.19}$$

provided the matrix $\boldsymbol{\Phi}$ is invertible. Consequently,

$$\mathbf{J} = \mathbf{V} \cdot \log{(\boldsymbol{\Lambda})} \cdot \mathbf{V}^{-1} \frac{1}{\Delta t} \tag{3.20}$$

where $\mathbf{V}$ is a matrix of right eigenvectors of $\mathbf{F}$ and $\boldsymbol{\Lambda}$ is a diagonal matrix of corresponding eigenvalues. The $\boldsymbol{\Phi}$ matrix corresponds to the coordinate transformation from the concentration space of eq (3.10) to the "canonical" space of eq (3.11).

For many chemical systems, some of the modes are too fast to be observed, and therefore the states that are actually observed are essentially confined to a p dimensional subspace, where $p < n$. In this case, eq (3.19) can no longer be used directly, since $\boldsymbol{\Phi}$ becomes ill-conditioned. The problem can be traced back to a degeneracy arising in the autoregression model (3.13). Partial extraction of information about $\mathbf{F}$ is still possible, however. We assume that the eigenvalues of $\mathbf{F} = e^{\mathbf{J}\Delta t}$ can be separated into two groups, $\boldsymbol{\rho_i}$ and $\boldsymbol{\rho_j}$, where $i = 1 \ldots p$ and $j = p + 1 \ldots n$, such that $|\boldsymbol{\rho_i}| \gg |\boldsymbol{\rho_j}|$. The first group corresponds to the slow modes, while the second group corresponds to the fast modes.

The autoregression model, identified from the motion in the slow subspace, will have a reduced order of p, and, correspondingly, matrices $\mathbf{L}$ and $\mathbf{C}$ will become $p \times p$. However, we can still perform n experiments with n perturbing species, obtaining the full rank $n \times n$ matrix $\mathbf{G}$ and $p \times n$ matrix $\mathbf{B}$. Equation (3.18) is still well defined in this case, but $\boldsymbol{\Phi}$ becomes a rectangular $p \times n$ matrix and does not provide a one-to-one correspondence between the canonical space and the concentration space. If we represent the canonical transition matrix $\mathbf{L}$ through its left eigenvectors $\mathbf{U}$, where $\mathbf{L} = \mathbf{U}^{-1} \cdot \boldsymbol{\Lambda} \cdot \mathbf{U}$, then we can obtain the left eigenvectors $\mathbf{W}$ for the slow modes by using the coordinate transformation provided by $\boldsymbol{\Phi}$:

$$\mathbf{W} = \mathbf{U} \cdot \boldsymbol{\Phi} \tag{3.21}$$

The left eigenvectors for the slower (measurable) modes are recovered by this method. If all of the modes are measurable, we recover the full set of left eigenvectors, and by inversion of the matrix of the left eigenvectors we obtain the full set of right eigenvectors. Knowing both the left and right eigenvectors, together with the corresponding eigenvalues, allows us as in eq (3.20) to determine the transition matrix for the discrete model and the corresponding Jacobian matrix for the continuous model. If m modes are so fast that their dynamics cannot be recovered, we assign zero eigenvalues to these modes (where zero eigenvalues of the transition matrix correspond to infinitely large negative eigenvalues of the Jacobian). We can still determine the left eigenvectors for the slower modes, as previously described. However, we do not obtain a complete set and thus we cannot use matrix inversion to obtain the unique right eigenvectors.

3.7 References

[1] J. J. Tyson, "Classification of Instabilities in Chemical Reaction Systems," *J. Chem. Phys.* **62**, 1010-1015 (1975).

[2] M. Eiswirth, A. Freund, and J. Ross, "Mechanistic Classification of Chemical Oscillators and the Role of Species," in I. Prigonine and S. A. Rice, eds., *Adv. Chem.Phys.* **80**, (Wiley, New York 1991).

[3] F. Hynne, P. G. Sørensen, and T. Møller, "Complete Optimization of Models of the Belousov-Zhabotinsy Reaction at Hopf Bifurcation," *J. Chem. Phys.* **98**, 211-230 (1993).

[4] T. Chevalier, I. Screiber, and J. Ross, "Toward a Ststematic Determination of Complex Reaction Mechanizms," *J. Phys. Chem.* **97**, 6776-6787 (1993).

[5] A. Arkin and J. Ross, "Statistical Construction of Chemical Reaction Mechanisms from Measured Time-Series," *J. Phys. Chem.* **99**, 970-979 (1995).

[6] R. Díaz-Sierra, J. B. Lozano, and V. Fairén, "Deduction of Chemical Mechanisms from the Linear Responce around Steady State," *J. Phys. Chem.* **103**, 337-343 (1999).

[7] P. G. Sørensen and F. Hynne, "Amplitudes and Phases of Small-Amplitude Belousov-Zhabotinsii Oscillations Derived from Quenching Experiments," *J. Phys. Chem.* **93**, 5467-5474 (1989).

[8] F. Hynne, P. G. Sørensen, and K. Nielsen, "Quenching of Chemical Oscillations: General Theory," *J. Chem. Phys.* **92**, 1747-1757 (1990).

[9] V. Vukojević, P. G. Sørensen, and F. Hynne, "Quenching Analysis of the Briggs-Rauscher Reaction," *J. Phys. Chem.* **97**, 4091-4100 (1993).

[10] G. C. Goodwin and K. S. Sin, *Adaptive Filtering, Prediction, and Control,* (Prentice Hall: Englewood Cliffs, N.J., 1984).

[11] V. Petrov, E. Mihaliuk, S. K. Scott, and K. Showalter, "Stabilizing and Characterizing Unstable States in High-Dimensional Systems from Time Series," *Phys. Rev. E* **51**, 3988-3996 (1995).

[12] W. Press, B. Flannery, S. Teukolsky, and W. Vetterling, *Numerical Recipes: The Art of Scientific Computing,* (Cambridge University Press: Cambridge, 1986).

[13] G. Golub and C. Van Loan, *Matrix Computations,* (Johns Hopkins University Press: Baltimore, 1983).

[14] R. J. Field and R. M. Noyes, "Oscillations in Chemical Systems. IV. Limit Cycle Behaviour in a Model of a Real Chemical Reaction," *J. Chem. Phys.* **60**, 1877-1884 (1974).

[15] A. N. Zaikin and A. M. Zhabotinsky, "Concentration Wave Propagation in Two-dimensional Liquid-phase Self-oscillatory Reaction," *Nature* **225**, 535-537 (1970).

[16] K. Nielsen, F. Hynne, and P. G. Sørensen, "Hopf Bifurcation in Chemical Kinetics," *J.Chem. Phys.* **94**, 1020-1029 (1991).

[17] K. Nielsen, *Hopf Bifurcation in Oscillatory Chemical Reactions,* Ph.D. Thesis, (University of Copenhagen, 1991).

[18] The offset parameter b_0, previously introduced for convenience in the autoregression calculation, is not used in determining the Jacobian matrix and is therefore omitted.

[19] M. Eigen, L. De Mayer, in A.Weissberger, ed., *Techniques of Organic Chemistry,* vol. **8/2**, (Interscience, New York, 1963).

[20] C. F. Bernasconi, *Relaxation Kinetics,* (Academic Press: New York, 1976).

[21] R. C. K. Lee, *Optimal Estimation, Identification, and Control,* Res. Monograph **26**, (M.I.T. Press: Cambridge Massachusetts, 1964).

Chapter 4

Locally-Linear Prediction
of the Chaotic Flame Front

4.1 Introduction

Forecasting enables one to choose actions with more desirable outcomes. The fundamental prediction and planning mechanisms, perfected by the millennia of evolution, are argued to be among the core building blocks of intelligence.[1,2] It is something so basic and automatic that we do not even notice its very existance. The awe at the efficiency of these intricate capabilities grows even greater when one attempts to replicate them in a scientific or technological domain. Understanding the general principles that govern practical high-dimensional predictors is thus not only of considerable applied interest but also an enquiry into our own nature. It is worth pointing out that low-dimensional methods typically do not scale up for high-dimensional systems, and there are non-trivial and often overlooked pathologies specific for high-dimensional spaces.[3] This makes modeling and prediction of complex systems and signals quite challenging and requires methods specifically designed for the purpose. In this chapter we present an approach to the prediction of high-dimensional systems based upon a locally-linear map in the embedding space. We demonstrate an efficient realization of such methods using a modified Relative Neighborhood mesh with local restructuring. The method is applied to prediction of both the future behavior and behavior at distant spatial locations of the chaotic flame front using time series of the

front position at a single spatial location.

Many real world systems can be adequately described on the macroscopic level by the deterministic laws, in which case the future of the system can, in principle, be derived from its current state. However, the erasure of the state information, inherent in chaotic systems sets practical limits to the actual extent of the possible prediction. The bounds on the extent and precision of the prediction are set by the local Lyapunov exponents of the system, available signal to noise ratio in the measurements and the prediction method.[4] Considering the properties of the system and the measuring process fixed, we focus on the prediction method itself.

It can be demonstrated[5,6] that the problem of prediction of a nonlinear dynamical system, whenever feasible, is equivalent to the problem of approximating the relationship between an input vector describing the current state of the system in a suitably chosen space and the desired output or future state. Intuitively, this follows from the embedding result by Takens[7,8] and determinism of the system dynamics. The choice of the approximation method depends upon the availability of additional knowledge about the system.

If a specific class of functional forms is preferred as a description of the empirical mapping, a small number of parameters are adjusted to tailor a function for the best fit to the empirical data. This constitutes the essence of the parametric methods and the obvious prior assumption here is the choice of the class of functions used to describe the input-output relationship. The most important feature is the small number of degrees of freedom (parameters) that are adjusted, thus making the task possible without requiring an infinite amount of input data for high-dimensional systems.

If there is no reason to favor any particular functional description, data-based models are used. In this case, the *a priori* assumptions are less obvious; however such assumptions are required for any method. Since there is an infinite number of functions (curved surfaces) passing through a given set of points in space, there must be a criterion to select one of these infinite possibilities. The assumption hidden in most of the data-based methods comes in the form of smoothness constraints on the input-output function.

Neural networks are often used for data-based approximation.[9] The structure of

the network and the learning algorithm determine its properties as a function approximator. It has been shown[10,11] that several of the most popular structures can be derived in a unified way using slightly different smoothness constraints. Unfortunately, there is a considerable burden in using neural-network approximators: their implementations are typically slow, they are hard to analyze with respect to performance problems, and the extraction of information from the trained network is difficult.

A locally linear approximation of the input-output function can be used to overcome these deficiencies. Generally suboptimal, this approach has a number of practical advantages over more advanced methods. As we show further, it lends to an efficient implementation, is robust in the presence of moderate noise, and its intrinsic organization of data is naturally suitable for analysis of the phase space of the system. The simplicity of the method allows one to analyze general problems associated with high-dimensional modeling with minimal artifacts from the method itself. In the rest of this chapter, we describe the details of the method and study its properties and performance in a realistic situation, using as a model a high-dimensional, spatially extended chaotic system such as the Kuramoto-Sivashinsky (KS) equation.

4.2   Method

Linear models and associated identification, prediction and control methods[12] are extremely well developed and offer a number of desirable properties. It has been demonstrated[13] that in a vicinity of a steady state or a periodic orbit, many nonlinear systems can be treated by these methods. However, one can not expect a linear approach to be successful beyond these restricted domains. The flows in the phase space of practically any nonlinear system, including the KS model presented here, have non-trivial topology and it is not possible to represent them by a single linear model. The attempt to extend the range of applicability, while preserving the simplicity of linear models leads to the introduction of the locally-linear method. This approach introduces a multitude of linear models, each representing one small region of the phase space. Thus, at any given moment in time, a single linear model is used, which preserves its amenability to the analysis by the methods generally applicable

to the linear models. At the same time, as the system traverses the phase space, the model used is constantly changing, allowing for unrestricted topological complexity of the phase space, regardless of the effective dimensionality of the local model.

The following steps are required to construct and use a locally-linear model:

1. Employ time-delay embedding to convert the scalar measurements into vectors. Time delay space of sufficient dimensionality is isomorphous to other representations of the phase space, such as one in terms of individual system variables.[7,8]

2. Search for the trajectories in the past evolution of the system that passed close to the current point.

3. Carry out linear regression on the flow in the neighborhood of the current point using the neighboring trajectories from the history of the system.

4. Predict future trajectory based upon the known current position and the calculated flow.

While steps 1,3 and 4 are straightforward to implement efficiently, there is a difficulty with step 2, due to the large amounts of data necessary for an adequate representation of the flows throughout the phase space. Several alternatives exist for its implementation providing different tradeoffs between computational time and the memory required to store auxiliary data.

A naive implementation of the method involves a straightforward linear search through all of the available history, represented as a simple sequence of vectors. This is a serious limitation to the utility of the method, since the time required to search through the history is proportional to $nm \log(m)$, where n is the dimensionality of the data and m is the total number of vectors in the history.[14]

It seems that it should be possible to partition the data using one of the known methods based on the search trees to speed up the retrieval of the neighbors, as suggested by Farmer and Sidorovich.[15] The performance gains, however, cannot be realized in case of high dimensionality of the space. This is a non-trivial observation, which is better understood if one considers the relationship between the Euclidean distance between two points in high dimensional space and the distance between these

points along any single direction. For higher dimensional spaces, there is less and less correlation between the two, and due to this fact, partitioning becomes increasingly inefficient. Practically all of the tree must be traversed during the search, which results in very poor performance, considering the intrinsic complexity of the algorithm. A similar observation is presented by Kanerva[3] for other problems involving high-dimensional search spaces.

The alternative approach is based upon explicitly maintaining the references from the trajectories to their neighbors. However, the high dimensionality implies an exponentially high number of neighbors. To overcome this problem we do not explicitly store all of the neighborhood relationships, but only a small number of them, satisfying certain properties. We use local a modification of the relative neighborhood mesh (RNM).[16] RNM is defined as a linked set of nodes with the nodes i and j connected if the distance ij is smaller than either ik or jk for all k. The remaining relations are then efficiently computed by iterative closure of explicit neighborhood relationships.

What makes this approach particularly attractive is the observation that the set of neighboring trajectories identified from the history changes relatively slowly during the evolution of a chosen target trajectory. Maintaining this set requires the addition and removal of only a small number of members. The actual procedure is based upon iterating through the previous set of neighbors of the target trajectory, reevaluating their current membership, and evaluating the distances to the neighbors of the current neighbors as candidates for the inclusion into the set of neighbors. Those of the accessed trajectories that satisfy the given neighborhood criterion are retained as the new set of neighbors.

The running time of this procedure depends only on the number of the current neighbors, as opposed to the total size of the data set. It seems that it would be necessary to conduct an extensive search for neighbors during the construction phase of the required mesh structure. This can be avoided, however, by an incremental approach. Due to the design of the method, the neighborhood relations are known at any moment for the vicinity of the current location. Restructuring this small domain to include new data at run time is sufficient to maintain an efficient mesh.

The mesh created by a long sequence of such local restructurings contains only

a small number of artifacts violating the neighborhood criterion on the global scale. Furthermore, the number of the artifacts appears to reach a limit and does not grow significantly in the course of further remeshing. There is also no noticeable performance difference between incrementally constructed mesh and one obtained by application of the globally acting method. Since the time required to construct a local neighborhood mesh grows at least cubically with the number of vertices being considered, a small size of the domain being restructured at each step is the key factor that enables "growing" the mesh at practically a constant rate, regardless of the total amount of data collected.

Finally we note that phase space is isomorphous to the time-delay space, and we therefore use time delay coordinates. However, in general, these spaces have different metrics, which makes spherical neighborhoods in one space to correspond to ellipsoids in the other. Thus when speaking in terms of the phase space we must be cautious. We do not know how this complication affects the performance of the method, which is an important practical issue, especially in the presence of noise in the data. It is likely that our method is suboptimal in this respect.

4.3   Kuramoto-Sivashinsky Equation

We use the Kuramoto-Sivashinsky (KS) equation[17,18] as an example of a spatially extended system exhibiting complex chaotic behavior, which cannot be adequately described or predicted by previously employed methods.[13] Figures 4.1 and 4.2 illustrate the complexity of the dynamical behavior exhibited by the KS equation.

The Kuramoto-Sivashinsky equation is one of the simplest nonlinear partial differential equations for modeling spatiotemporal chaos. It has been found to mimic the dynamical behavior of many different physical systems but is most often used to model the spatiotemporal evolution of 2D flame fronts.[17] The governing equation for the contour of the front has the form

$$\frac{\partial \psi}{\partial t} + \left(\frac{\partial \psi}{\partial z}\right)^2 + \frac{\partial^2 \psi}{\partial z^2} + \frac{\partial^4 \psi}{\partial z^4} = 0. \tag{4.1}$$

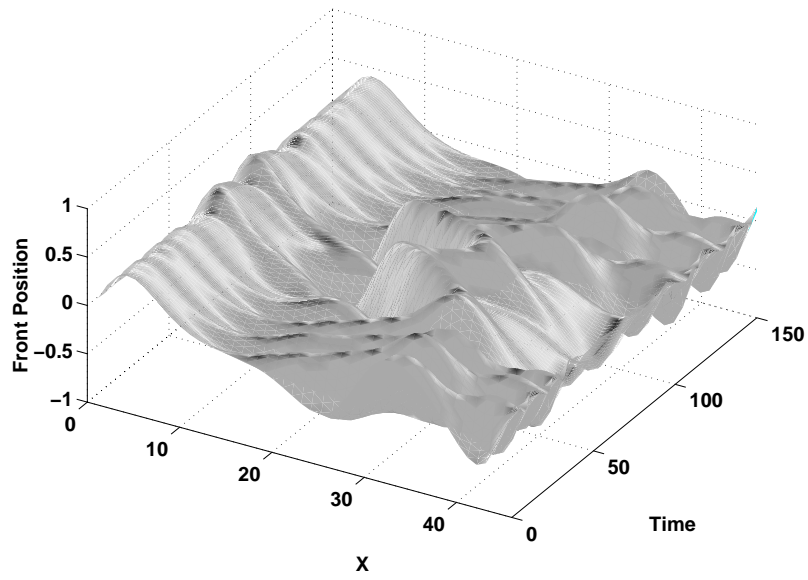The width L of the front is the only parameter in the system. Its value determines

Figure 4.1: A 3D space-time plot representing the oscillating wave front. The left axis is time t, the right axis is the position x, along the front, and the vertical axis is the variable.
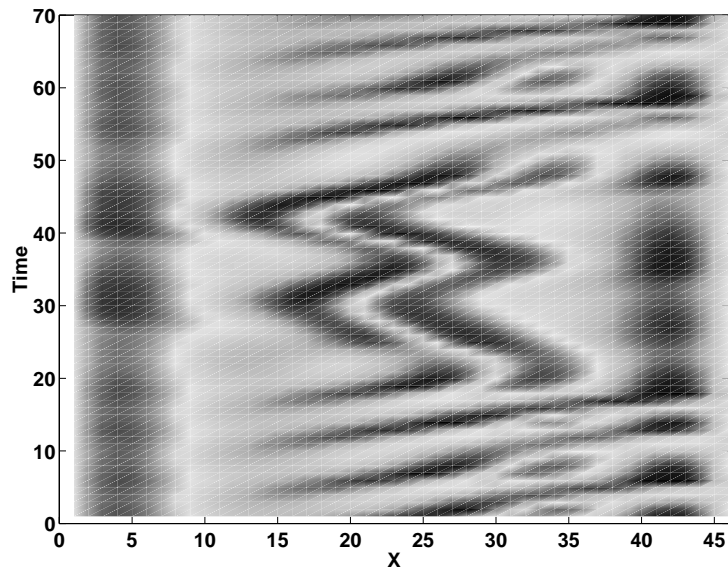
Figure 4.2: A space-time plot of the oscillations, with values of the represented by gray scale levels. The abscissa is the position x along the front, the ordinate is the time. Note the complex character of the oscillations observed in figures one and two.
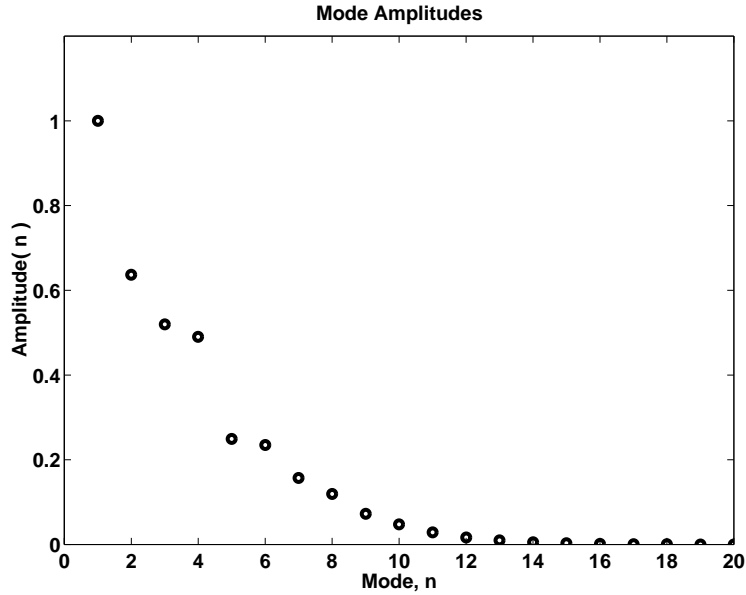
Figure 4.3: Singular values of the oscillation modes of the KS equation.

the dynamical properties of the front, which range from a simple steady state for narrow fronts to fully developed turbulence for wider ones.[19] We choose a moderate width $L = 18.3$ where a 2-cell regime is competing with a 3-cell regime resulting in chaotic oscillations.

Normal mode analysis yields the oscillatory modes having the largest contribution to the spatiotemporal shape of the front. The shape of the oscillating wave front can be represented as a sum of basic profiles. SVD[20] is used to construct a set of orthogonal modes, which are ordered by the decreasing variation of their coefficient in the sum. The singular value is this variation, i.e., the first mode accounts for most of the wave front dynamics, the next one is less important but it accounts for most of the dynamics not accounted for by the first mode, and so on. Note that the shape of the wave front can be described with an error of less than 0.1 by the first twelve modes.
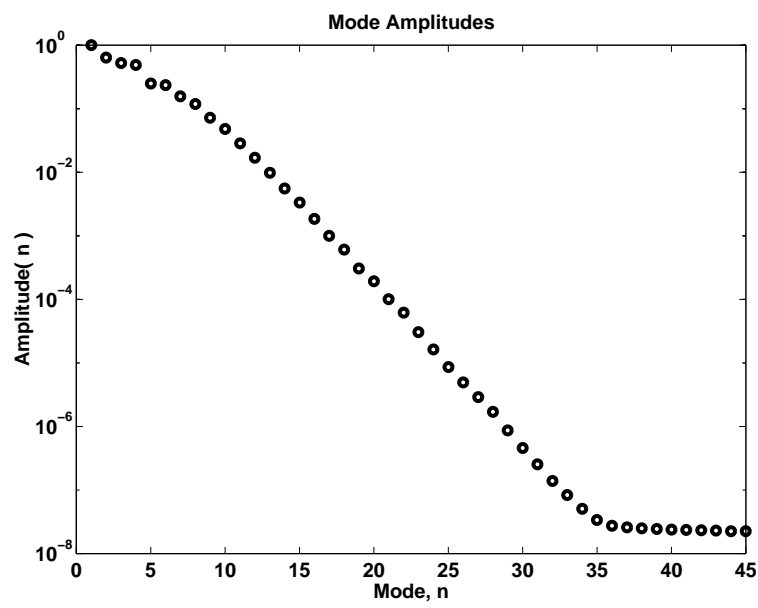
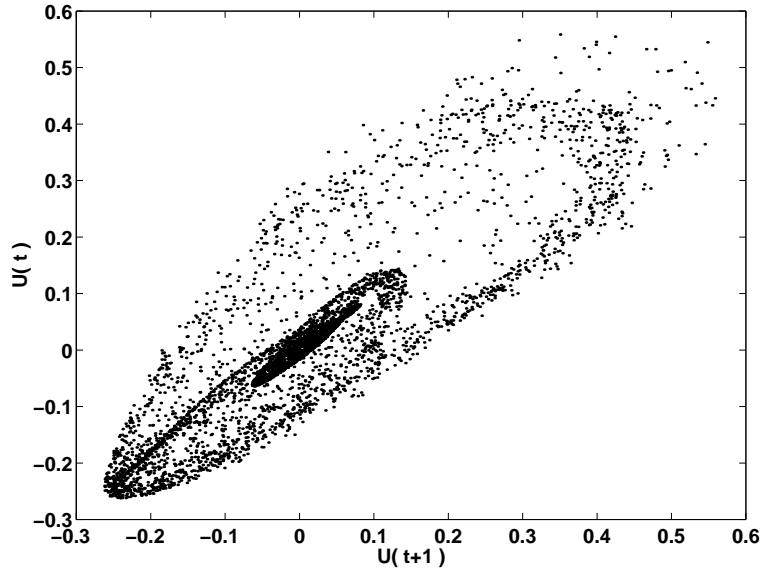**Figure 4.4:** Logarithmic plot of singular values in Figure 4.3.

Figure 4.5: Time delay plot of ξ at a fixed spatial location $x = 10$. The abscissa is $\xi(t)$, while the ordinate is $\xi(t + \Delta t)$, where $\Delta t = 2$.

Figure 4.4 is identical to the Figure 4.3 except that a logarithmic scale is used for the singular values, to better represent their full range of values. This plot is a manifestation of the true complexity of the wave front dynamics: Even though the impact of the higher order modes decreases, the number of modes necessary to exactly represent the dynamics is very large. It has been shown, however, that this number is finite.[19] The dominant influence of a few largest modes is also evident from the following scatter plots.

The plot in Figure 4.5 demonstrates a complex but non-random dependence of the successive values, which is a key requirement for the predictability of the time series. Generally, predictability can be related to the mutual information between current readings and vectors of the past readings of the variable. The time delay plot thus illustrates the simplest one-dimensional case of this dependency. The clustering

63

of trajectories clearly indicates the possibility of their prediction, which can only improve with the use of higher order dependancies.

Figure 4.6 shows that there is a certain degree of interdependence between the values at the spatially distant locations, as expected from the mode analysis. If all of the dynamics were explainable by a single spatial mode, then values at different locations would be proportional to each other, even with chaotic time series. The varying influence of several different modes causes the spatial complexity. If n modes provide a complete description of the front shape, it is then necessary to perform measurements at n non-degenerate locations to exactly predict the position everywhere in the front. The low dimensionality of the system dynamics on a coarse scale is evident from the high degree of localization present in the above plots.

4.4   Results

As described in Chapter 2, the dynamics very close to a steady state of a fairly wide KS front ($L = 35.0$) can be adequately described by a linear model. This description, however, is valid for only a very small region surrounding the steady state. In the case of autonomous chaotic oscillations, excursions of the system cover a considerable fraction of the available phase space, and it is therefore not surprising that a single linear model does not provide a faithful description to these flows in the phase space. Figures 4.7 and 4.9 illustrate the failure of the linear models to capture the dynamics of the system. In Figure 4.7 we demonstrate the prediction of the future values at a given spatial location. In Figure 4.9 we predict simultaneous, but spatially distant values. Note the poor performance of the linear predictors, regardless of order. A linear predictor is the simplest possible model, which is also easy to fit to the available data in a way that minimizes the RMS prediction error.

For the nonlinear method, ten thousand time instances were collected to represent the history of the system. Another five hundred instances were then both simulated and predicted using the data-based locally linear SDM. We now present the RMS prediction error as a function of temporal and spatial extent of the prediction interval.

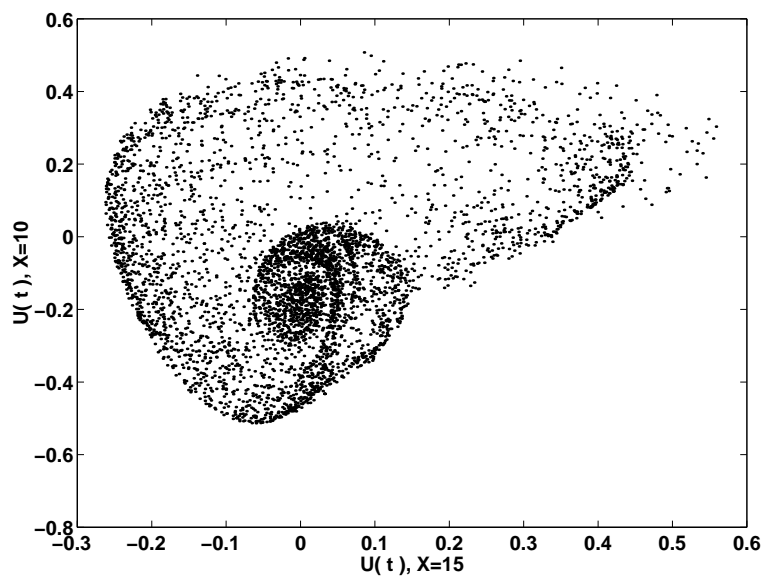Note the dramatic improvement of the prediction quality at the expense of a much

Figure 4.6: Scatter plot of values at one spatial location ($x = 10$) as related to the simultaneous readings at another location ($x = 30$).
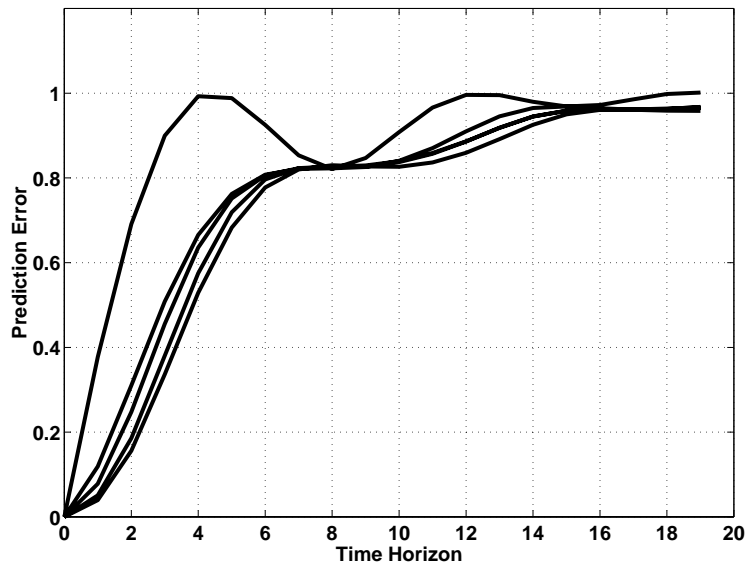
Figure 4.7: RMS prediction error of linear predictors of different order ($n = 1..5$) as a function of iterates into the future predicted. The value in the future is calculated as a linear combination (with constant coefficients) of a sequence of n values ending with the current reading.
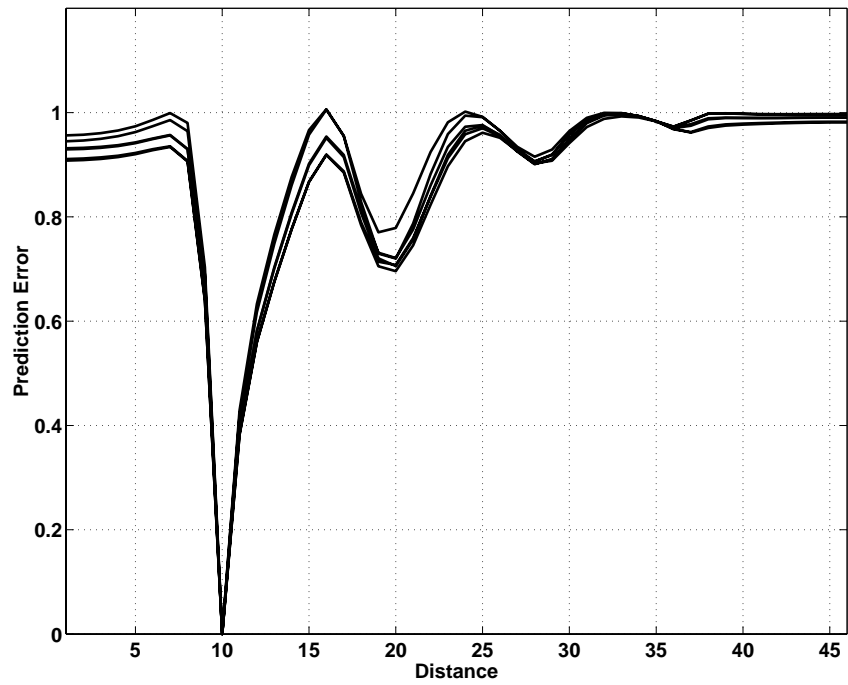
Figure 4.8: RMS prediction error of linear predictors of order $n = 1..5$ as a function of the predicted point location. The vector of the measurements is taken at the spatial location $(x = 10)$.
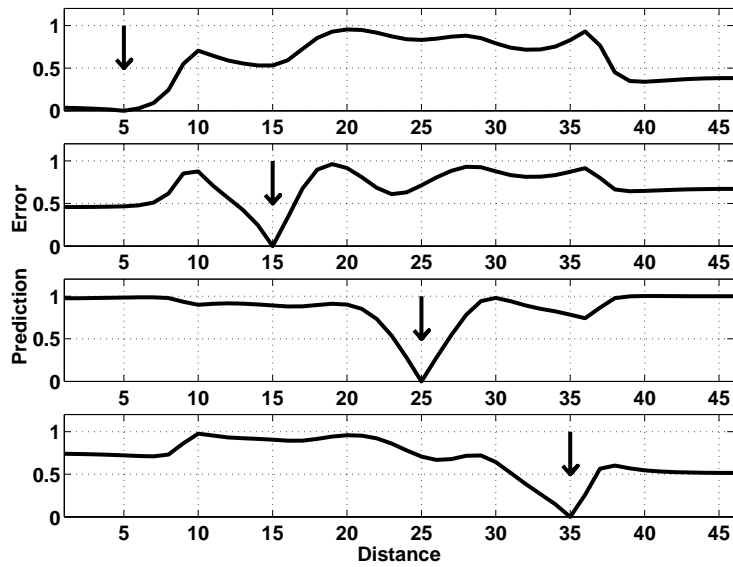
Figure 4.9: RMS prediction error of linear predictors of order $n = 5$ as a function of the distance between the predicted point and the point where the vector of the measurements is taken. The different plots demonstrate the prediction using measurements at four different spatial locations.

larger amount of data constituting the predictor. In the linear case $n = 5$, there are just five constant coefficients. In the $n = 5$ SD predictor, we maintain the history consisting of 10,000 measurements to recompute the five coefficients at every point in the phase space whenever necessary. In fact, we use the history as the sampled representation of the flow in the phase space and compute a linear approximation of this flow at every point were the prediction is attempted. This leaves the global structure of the flow completely unconstrained by the model and it is defined solely by the experimental data available. Trying to approximate this structure with a single linear predictor in many cases is not possible, obviously so if there is more than one singularity in the phase space.

4.5   Conclusion

An efficient implementation of the multidimensional prediction method has been presented. The method was developed as a generalization of the previous work on linear methods[13] and shows improved performance and wider range of applicability. It offers a concrete mechanism allowing the application of recently popularized nonlinear control and prediction principles.[6] The approach falls into the general class of State Dependent Models.[21] However, known previously published examples of SDM have been limited to low-dimensional cases,[22] whereas we have been able to demonstrate prediction of a spatially extended chaotic system. We recently became aware of the work by Farmer and Sidorovich,[15,23] proposing a method substantially similar to ours. However, in their work the authors propose a different approach to the implementation of the idea. As we have already discussed, the efficiency of the high-dimensional implementation is of singular importance for the practical applicability of the method. We consider the development of the mesh-based method to be an improvement over the implementations discussed by Farmer and Sidorovich.[15] The performance of the method was verified by its application to the temporal and spatial prediction of the Kuramoto-Sivashinsky wave front.
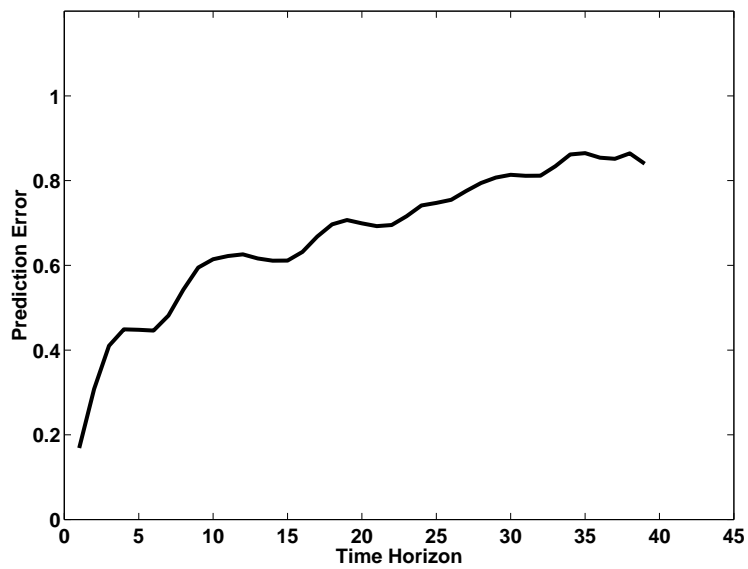
Figure 4.10: RMS prediction error for SD predictors of different order (a) $n = 4$, (b) $n = 5$ as a function of iterates into the future predicted. Analogous to Figure 4.7, but with state-dependent predictor.
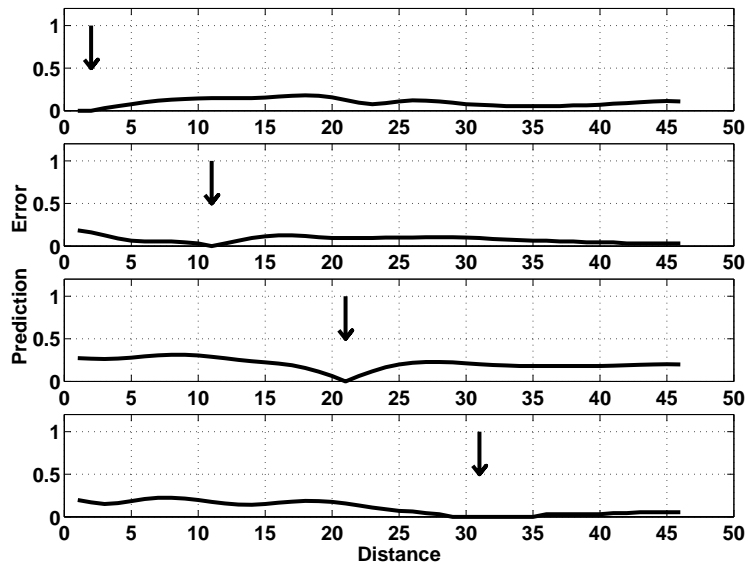
Figure 4.11: RMS prediction error of SD predictors of order $n = 5$ as a function of distance between the predicted point and the point where the vector of the measurements is taken. The different plots demonstrate prediction using measurements at four different spatial locations. Analogous to Figure 4.9, but using state-dependent predictor.

4.6   References

[1] W. H. Calvin, "The unitary hypothesis: A common neural circuitry for novel manipulations, language, plan-ahead, and throwing?" in K. R. Gibson and T. Ingold, eds., *Tools, Language, and Cognition in Human Evolution,* (Cambridge University Press, 1993).

[2] D. Rumelhart, J.McClelland, and the PDP research group, *Parallel Distributed Processing,* (MIT Press 1986).

[3] P. Kanerva, *Sparse Distributed Memory,* (MIT Press 1988).

[4] H. D. I. Abarbanel, R. Brown, and J. B. Kadtke, "Prediction in Chaotic Nonlinear Systems: Methods for Time Series with Broadband Fourier Spectra," *Phys. Rev. A* **41**, 1782-1807, (1990).

[5] A. Isidori, *Nonlinear Control Systems,* (Springer-Verlag 1995).

[6] V. Petrov, and K. Showalter, "Nonlinear Control of Dynamical Systems from Time Series," *Phys. Rev. Lett.* **76**, 3312-3315 (1996); V. Petrov, et al., "Nonlinear Control of Remote Unstable States in a Liquid Bridge Experiment," *Phys. Rev. Lett.* **77**, 3779 (1996).

[7] F. Takens, "Detecting Strange Attractors in Turbulence," in *Lecture Notes in Mathematics* **898**, (Springer-Verlag 1981).

[8] T. Sauer, J. A. Yorke, and M. Casdagli, "Embedology," *J. of Stat.Phys.* **65**, 579-616 (1991).

[9] I. Grabec, "Prediction of Chaotic Dynamical Phenomena by Neural Network," in B. Souĉek and the IRIS Group, *Dynamic, Genetic, and Chaotic Programming,* (John Wiley & Sons 1992).

[10] T. Poggio, and F. Girosi, "Learning Algorithms and Network Architectures," in T. Poggio and D. Glaser, eds., *Exploring Brain Functions: Models in Neuroscience,* (John Wiley and Sons 1993).

[11] F. Girosi, M. Johns, T. Poggio, "Regularization Theory and Neural Network Architectures," *Neural Computation* **7**, 219-269 (1995).

[12] R. C. K. Lee, *Optimal Estimation, Identification and Control,* Research Monograph 28, (MIT Press 1964).

[13] V. Petrov, E. Mihaliuk, S. K. Scott, and K. Showalter, "Stabilizing and Characterizing Unstable States in High-dimesional Systems from Time Series," *Phys. Rev. E* **51**, 3988-3996 (1995).

[14] N. Wirth, *Algorithms + Data Structures = Programs,* (Prentice-Hall 1976).

[15] J. Farmer, and J. Sidorovich, "Predicting Chaotic Time Series," *Phys. Rev. Lett.* **59**, 845-848 (1987); J. Farmer, and J. Sidorovich, "Exploiting Chaos to Predict the Future and Reduce Noise," in Y. Lee, ed.,*Evolution, learning and cognition,* (World Scientific Publishing 1988).

[16] S. Arya, D. Mount, *Algorithms for Fast Vector Quantization,* Technical Report CS-TR-3017, (University of Maryland Institute for Advanced Computer Studies (UMIACS) 1993).

[17] G. I. Sivashinsky, *Ann. Rev. Fluid Mech.* **15**, 179 (1983).

[18] Y. Kuramoto, "Instability and Turbulence of Wavefronts in Reaction-Diffusion Systems," *Progress of Theoretical Physics* **63**, 1885-1903 (1980).

[19] B. Nicolaenko, B. Scheurer, and R. Temam, "Some Global Properties of the Kuramoto-Sivashinsky Equations: Nonlinear Stability and Attractors," *Physica D* **16**, 155-183 (1985).

[20] G. H. Golub, and C. F. V. Loan, *Matrix Computations,* (Johns Hopkins Series in the Mathematical Sciences 1996).

[21] M. Pristley, *Non-linear and Non-stationary Time Series Analysis,* (Academic Press 1988).

[22] R. Chen, "Threshold Variable Selection in Open-Loop Threshold Autoregressive Models," *J. of Time Series Analysis* **16**, 461 (1993).

[23] T. Sauer, "Time Series Prediction by using Delay Coordinate Embedding," in E. Ott, T. Sauer, and J. A. Yorke, eds., *Coping with Chaos,* (John Wiley and Sons 1994).

Chapter 5

Entrainment with Hebbian Learning

5.1 Motivation

There is a great degree of unification in the biological world. The genes initiating the construction of an eye are very similar, in fact, interchangeable, in such different organisms as flies and mice.[1] Molecular biologists understand the fundamental mechanisms used to encode the body plan of a particular animal in its genome.[2] These mechanisms are also almost exactly preserved across the species. Similarly, the "control system" of any animal is made up of neurons, specialized cells with electrically active membranes that serve, depending upon their exact nature and state, as filters, amplifiers, oscillators and other building blocks of the nervous system. The complexity of the system arises not from the elements themselves, although they can individually perform highly complex functions,[3] but from the fact that there are myriads of them, with an even larger number of synaptic interconnections between them.[4] It is easy to demonstrate that such a large number of connections can not be independently specified by the genetic material available to the development process (see Appendix). Yet it is not known what mechanisms are responsible for the creation of this intricate organization. The unification mentioned above implies that there must be similarity in the mechanisms acting in different organisms, and insights gleaned from studies of insects, for instance, may apply to more complex species, like ourselves. The nervous systems of flies and cockroaches appear to be very similar in two individuals with identical genetic makup.[5] Thus, we can imagine the existance

of a constructive mechanism that unfolds the genetic program to indirectly guide the construction of a structure that can not possibly be directly specified. In higher organisms, even identical twins do not posses exactly the same brain structures. It is apparent that larger brains in more complex animals are at least as precisely specified as the nervous systems of simpler animals. Learning is concurrent with maturation in humans; however, for animals with shorter maturation period many non-trivial traits and instincts are inborn, i.e., genetically encoded and ready for use practically from birth. Thus, there are specific degrees of freedom in development that are not rigidly controlled by the genetic program, while genes continue to exercise precise control over the rest of the process. The questions arise: What degrees of freedom are most likely to be left out of the exact specification ? What happens to these degrees of freedom ?

One of the obvious differences between the cockroach nervous system and the human brain is the presence in the human brain of the large neural tracts, comprised of the millions of the axons. Another striking feature of the human brain is that it has a number of "patches" with topographic correspondence to the body parts or to other brain regions.[6,7] There is a significant economy for the encoding process to represent the connections on the level of tracts whenever possible, rather than individual axons. For such encoding to be possible, a mechanism must exist to provide "automatic" routing of the neurons within a given tract to their precise targets, without specifically designating these targets genetically. In the case of a neural tract connecting two isomorphic areas, such a mechanism appears to be most plausible, with the exact specifications replaced by certain self-organization mapping processes.

The visual system is one in which topographic maps are the rule rather than exception, although they are also common in auditory, somatosensory and motor areas.[8,9] Thus it seems natural to explore the available evidence for the tract routing processes in the visual system. This system starts with the eye, where optics produce a two-dimensional projection of the light from the environment onto $\approx 10^8$ photosensitive rod and cone cells. These sensors convert light quanta into electric potentials, and provide amplification with gain control to operate over the necessary dynamic range. The first rudimentary neural processing occurs in the eye, where six layers of neurons are incorporated in the retina, with $\approx 10^6$ axons emerging as the output to the brain

in the form of the optic nerve. The optic nerves from the two eyes cross at the optic chiasma, after which axons carrying information about the left half of the visual field from both eyes travel to the right and those carrying information about the right half travel to the left. They arrive at the dorsal Lateral Geniculate Nucleus (LGN), where information from magnacellular and parvacellular paths of the two eyes is segregated, processed, partially combined, and sent further, mostly to the V1 area of the primary visual cortex. The top layers of the LGN are in simple topographic correspondence with the retinas of the eyes. More complex patterning, due to the interaction of information from the two eyes necessary for binocular vision, exists in the lower layers and in most of the higher areas such as V1,V2,V4, MT *etc.* However, certain topographic features are preserved throughout the system.

It is known that the unscrambling of initially imperfect connections in the optic nerve requires visual input.[10,11] Pre-birth spontaneous wave activity in the retina, observed experimentally, has been suggested as the necessary input early in development.[12] The common explanation is that localized stimulation of the retina, together with coincidence-detecting mechanisms in the brain, are responsible for refining the connections. We develop this view further, showing how entrainment of wave-like activity in two imperfectly connected excitable layers – interacting with the mechanisms of plasticity, represented on the abstract level by Hebbian learning – results in improved connectivity. We suggest that establishing the precise mapping can be based upon an essential similarity in the dynamical properties of the source and the target layers.

A typical biological neuron receives input from as many as 10,000 other neurons. The strength of these connections is not fixed. The most general principle, due to Hebb,[13] postulates an increase in the connection efficacy proportional to the frequency of simultaneous excitation of both source and target neurons. Even this simplistic mechanism can account for a host of adaptation and learning phenomena. We point out the possibility of interaction between Hebbian learning and entrainment of activity in the target neural layer by the received input. Entrainment occurs when a dynamical system is subjected to a signal correlated with the dynamics of the system. This phenomenon can be considered as a generalized example of resonance. The effects of perturbations that are synchronous with the system evolution accumulate

and give rise to effects that are larger than those independent of the dynamics of the system. Thus, similar systems can exhibit considerable influence on each other even under relatively weak coupling, while random noise is much less effective in perturbing either of them. In this chapter we present the results of numerical simulations of a system combining the effects of Hebbian learning with entrainment and demonstrate the evolution of the connections between two similar neural layers under such conditions. The observed phenomena and conditions necessary for their occurrence closely resemble those present in the developing visual system and, possibly, other parts of the brain involving topological connections between layers of neurons. In the next section we present a simple example of an excitable medium and demonstrate entrainment of activity in this medium by a suitably chosen input.

5.2   Entrainment

Nerve cells transmit excitation as electrochemical pulses. Propagation of the pulse is due to electrical conductance along the cell axle (a small displacement of a large population of ions), with constant signal regeneration due to the transversal ionic currents (a large displacement of a comparatively small number of ions).[14]   The Hodgkin-Huxley model[14, 15] incorporates a detailed description of the ionic currents to produce a quantitatively accurate picture of the processes occurring at every point of the axon during the pulse propagation. The Fitzhugh-Nagumo equations are a simplification of this model:

$$
\begin{aligned}
\dot{u}_j &= \epsilon \cdot \left(v + u_j - \frac{u_j^3}{3} + z \right), \\
\dot{v}_j &= \frac{1}{\epsilon} \cdot \left(a - u_j - bv_j \right).
\end{aligned}
\tag{5.1}
$$

These equations with two variables can exhibit excitability and oscillations similar to those in real neurons, but they are intended only as a quantitative model of the action potential.[16] In these equations, u can be roughly identified with the voltage, and v accounts for the recovery processes; a, b, c, and z are the model parameters with typical values $a = 0.7$, $b = 0.8$, $c = 3$, $z = 0$; h is the time scale factor, which is superfluous unless used in a nonuniform spatially extended system (otherwise, $h = 1$
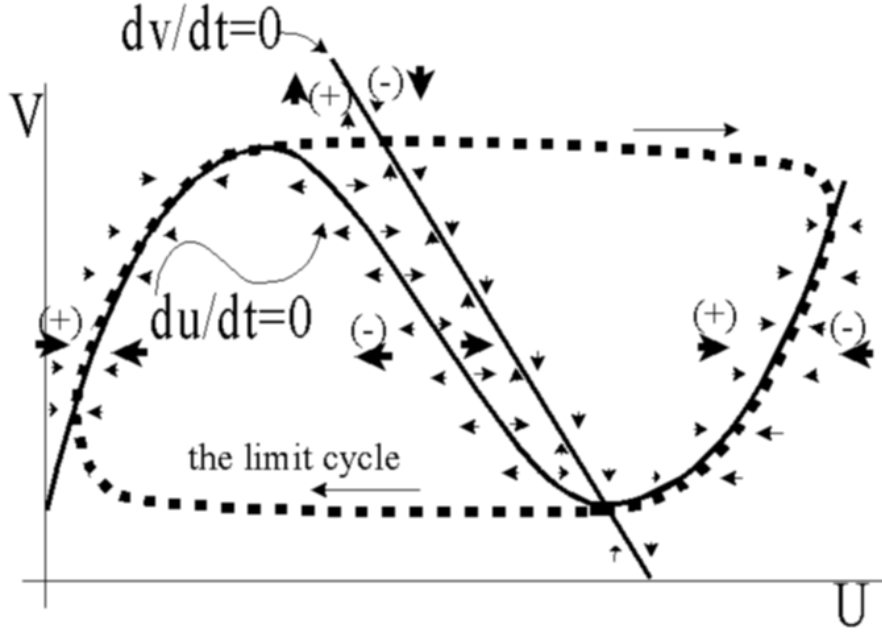
Figure 5.1: A phase portrait of the FitzHugh-Nagumo model.

). The values of the parameters determine the behavior in a manner that can be most easily understood upon analysis of the phase portrait of the system, as shown in Figure 5.1. The relative position of the nullclines, $y = (a - x)/b$ and $y = x^3/3 - x - z$, determines the character of the singular point at their intersection and thus the stability of the system, Figure 5.2. Changing the value of either a or z allows one to select an excitable, oscillatory or non-excitable regime.

Consider a one-dimensional excitable medium with point dynamics described by the Fitzhuh-Nagumo (FN) equations, and diffusive coupling between the elements:

$$
\begin{aligned}
\dot{u}_j &= \epsilon \cdot \left(v + u_j - \frac{u_j^3}{3} + z \right) + D_u \cdot (u_{j+1} + u_{j-1} - 2u_j), \\
\dot{v}_j &= \frac{1}{\epsilon} \cdot (a - u_j - bv_j).
\end{aligned}
\tag{5.2}
$$

This system of the ODEs can be considered as a discretization of the corresponding PDE, and as such is often used as a computational scheme for the numerical solution of the latter.[17] With the proper choice of parameters it possesses a traveling
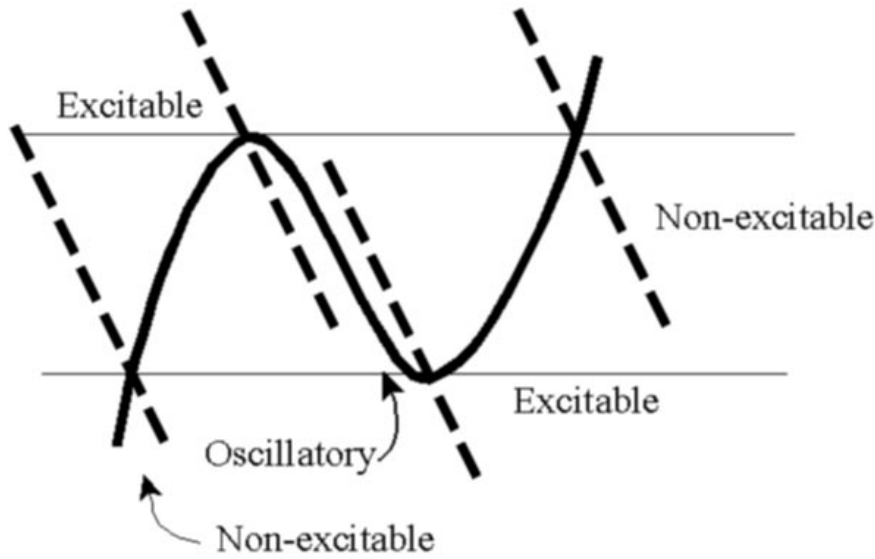
79

Figure 5.2: Stability of the FitzHugh-Nagumo model.

wave solution similar to that shown in Figure 5.4. In a biological context, the coupling between neighboring cells can be considered to be diffusive in the case of signal transmission by diffusion of extracellular factors, or in the case of the gap junction between cells. In case of synaptic contacts between the cells, however, transmission occurs only in one direction and can not be accurately described by the diffusive model. A comparison of the propagating wave of excitation in the array of synaptically coupled neurons, Figure 5.5, to one in a diffusively coupled array, Figure 5.4, reveals the qualitative similarity between the two. The diffusively coupled model will therefore be used throughout this work.

The propagation velocity depends upon the excitability of the system. Parameter z is introduced into the equation to provide a straightforward control of the excitation threshold, with more negative values corresponding to a more excitable system. The relationship between the value of the z parameter and the wave velocity is shown in Figure 5.6.

We use this model in its discrete interpretation to represent a layer of diffusively coupled neurons. The excitation projected onto this layer from outside is applied
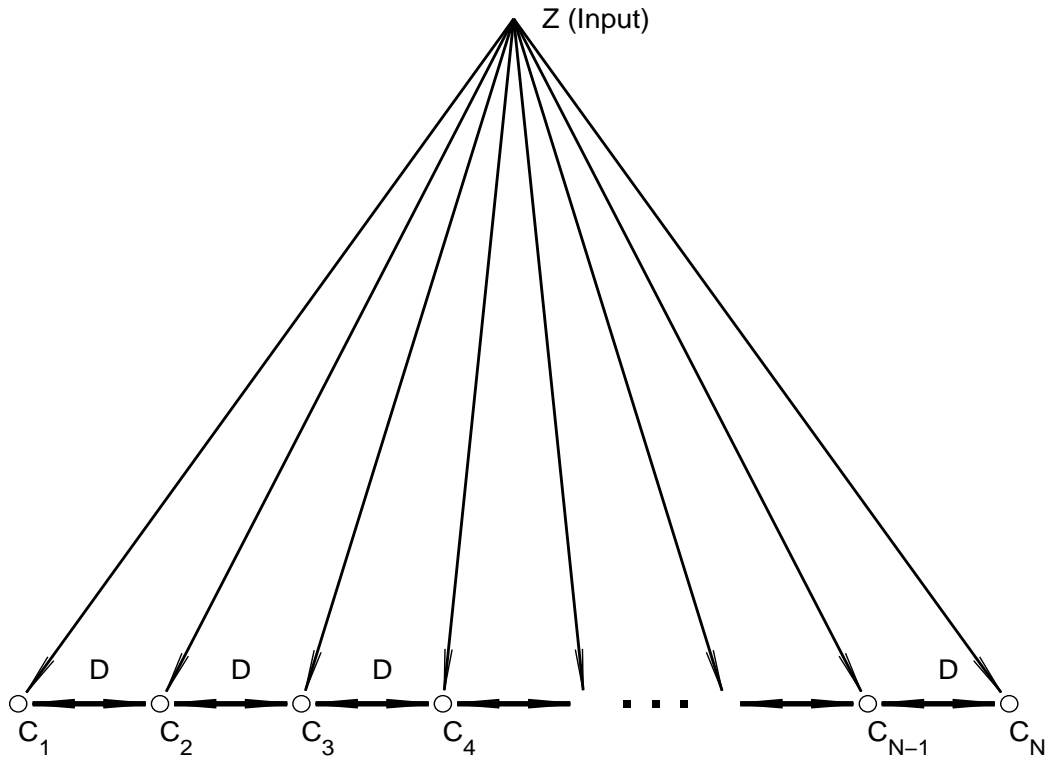
Figure 5.3: A one-dimensional array of excitable elements with diffusive coupling between neighbors and excitatory input.
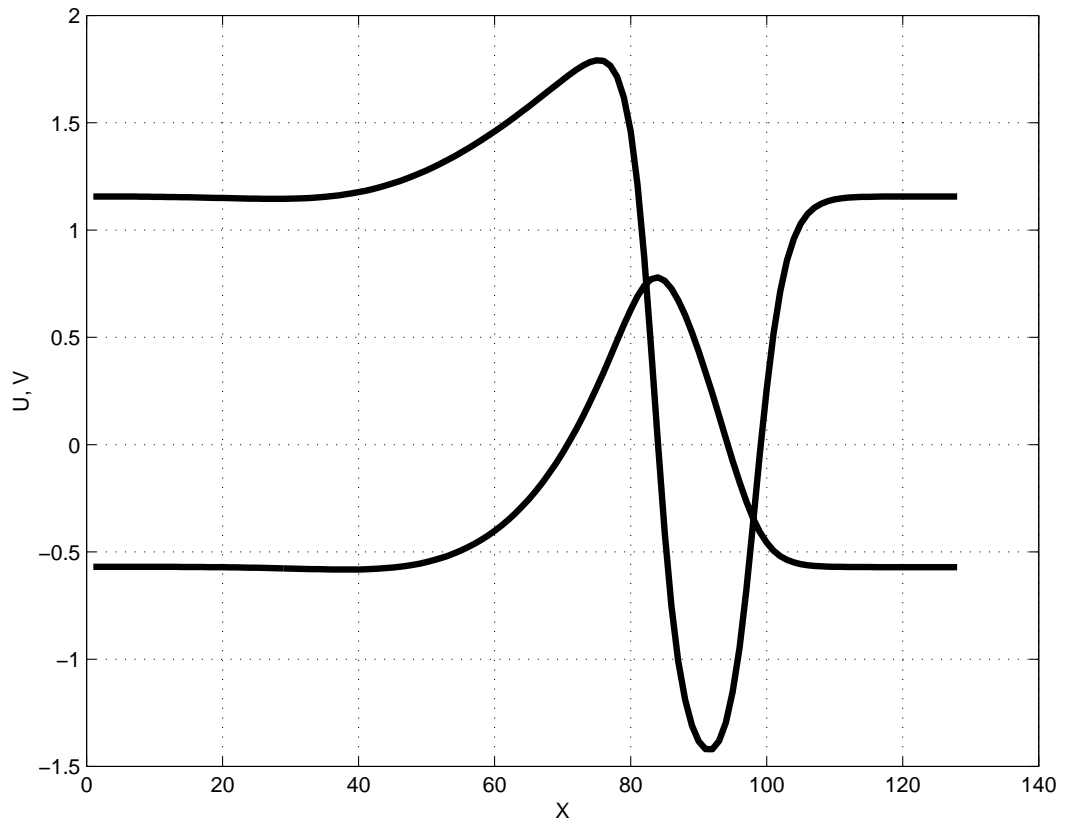
Figure 5.4: A typical excitation wave. The value of the variable u across the spatial array.
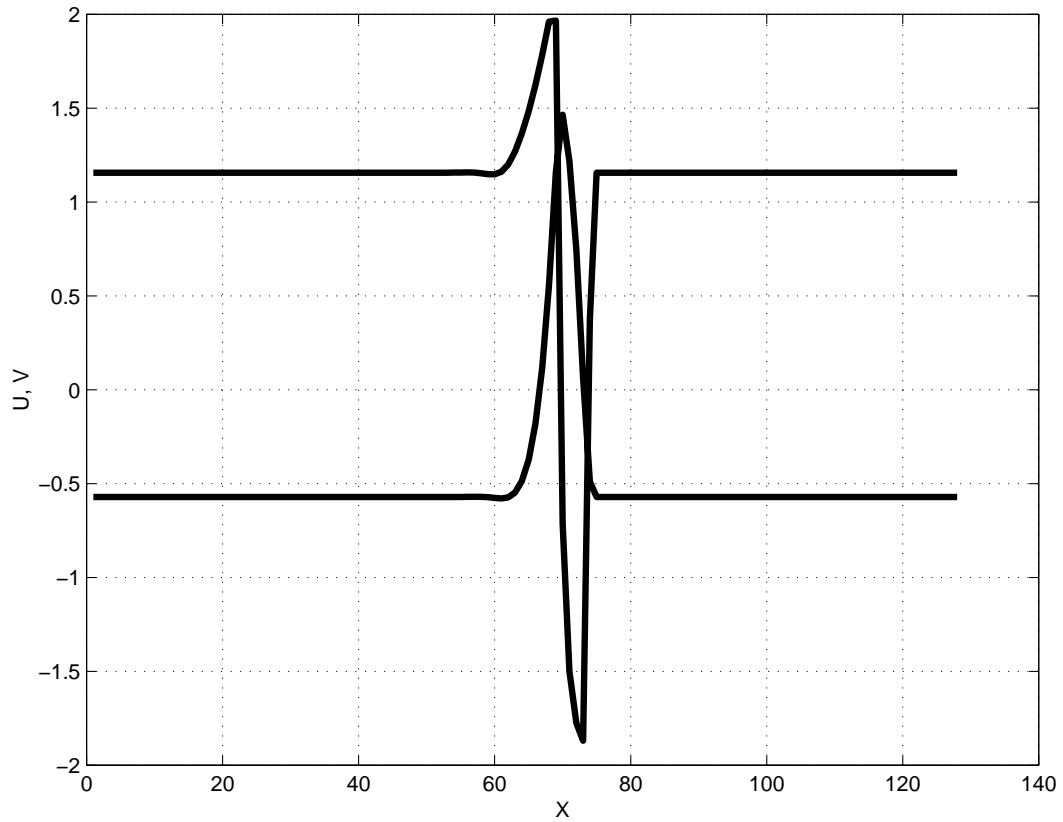
Figure 5.5: Similar to Figure 5.4, except coupling between the cells in one-way excitatory: if $u_i < 0$, add $\alpha = 0.01$ to z_{i+1} and z_{i-1}.
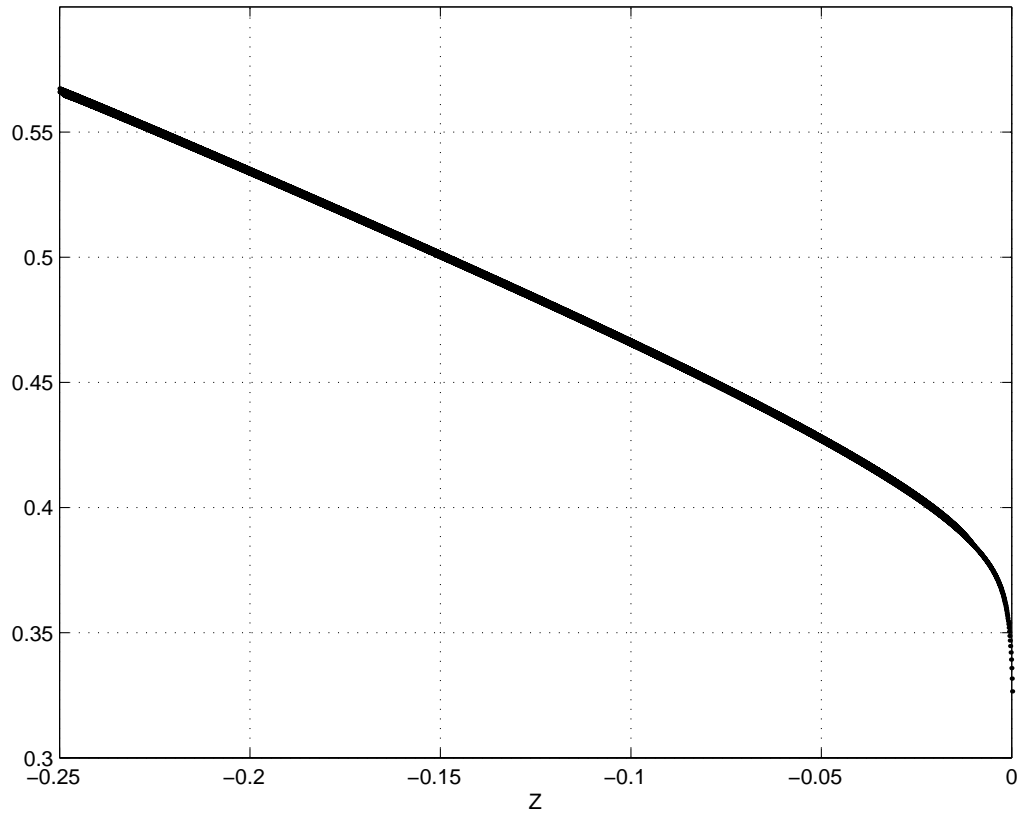
Figure 5.6: The velocity v of the wave as a function of the excitatory input. The value of constant uniform excitation is z.
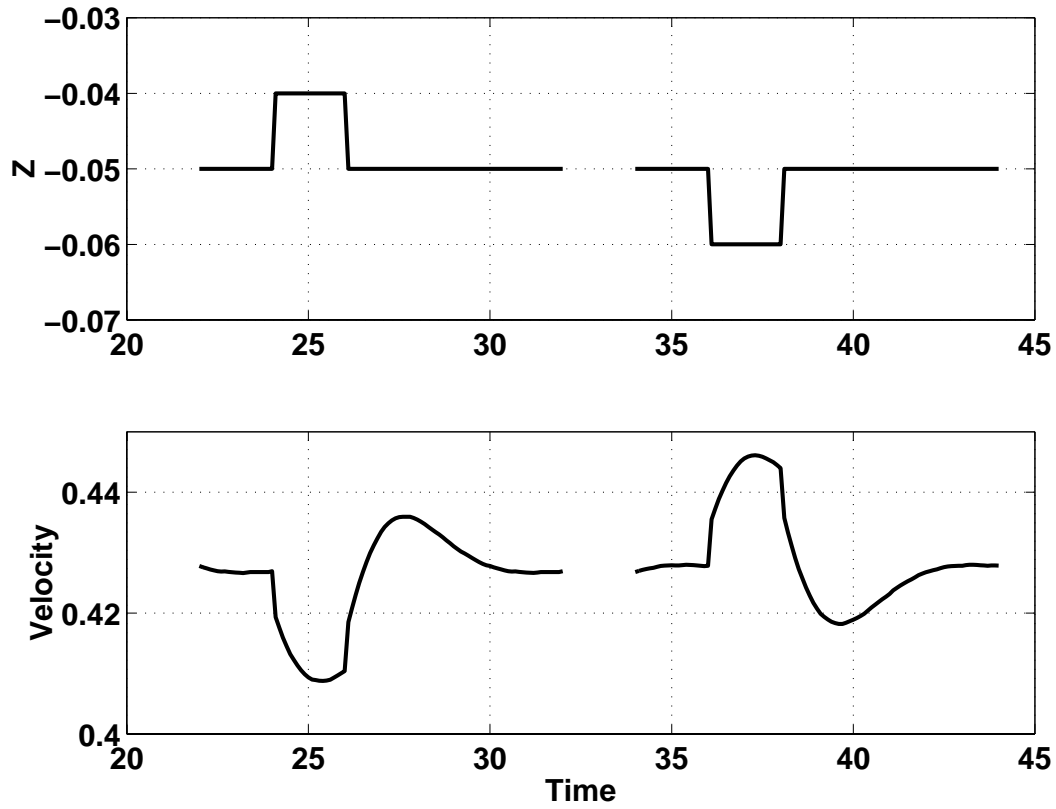
Figure 5.7: The change in the velocity of the wave v in response to a step-wise change in the uniform excitation z.
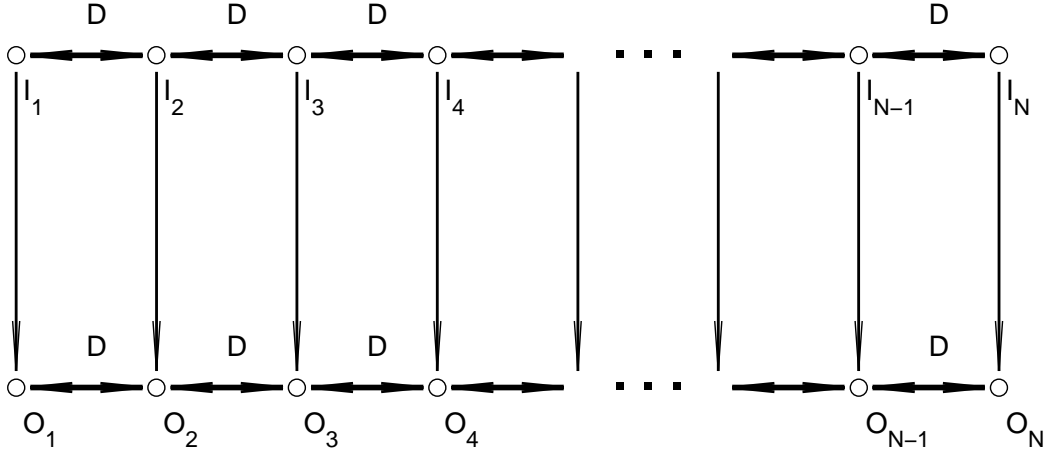
Figure 5.8: Two one-dimensional arrays of excitable elements with diffusive coupling between neighbors and one-to-one excitatory connections from the upper (source) to the lower (target) layers.

to the z parameter. The alternative choice is to make the input additive to the u variable. However, we believe that our method better accounts for the independence of the synaptic action upon the state of the target neuron. If the input reaching the layer is spatially uniform, it alters the threshold of excitation, and affects the propagation velocity in the manner that has been discussed above (Figure 5.6). The behavior of the system under inhomogeneous input conditions is of particular interest as a step toward understanding the interaction between coupled layers. It is natural to expect an acceleration of the wave at the regions receiving a stronger excitatory input and deceleration in the regions of weak input. This behavior is indeed present as illustrated by the Figure 5.7. It also shows that similar, although slightly more complicated, behavior is observed in response to the spatially uniform but time-varying excitation. The latter also demonstrates that the response of the system to a sequence of inputs cannot be simply predicted from the responses to the individual stationary patterns. This complication prevents drawing conclusions from these examples for the case of arbitrary inputs. However, the above results help in the interpretation of more complicated situations. Consider two identical layers, each similar to those described above. We will let the top (source) layer project one-way, excitatory connections to the bottom (target) layer in a one-to-one fashion (Figure 5.8):

$$z_i^t = z^t + \epsilon, \; u_i^s \geq threshold,$$
$$z_i^t = z^t, \qquad u_i^s < threshold. \tag{5.3}$$

If a wave is initiated in the top layer, it will project onto the bottom layer as a localized excitation, moving synchronously with the top layer wave. Intuitively, a wave in the target layer will move faster when coincident with the excitation and slower otherwise. This can be confirmed by arranging slightly different parameters for the layers ($z_{bottom} = z_{top} + \delta$) to make the autowave in the bottom layer move slightly slower than the one in the top layer and allowing them to interact in this adiabatic condition ($vs \approx vt$, since δ is small) when passing one another. Measurements of instantaneous wave velocity in the target layer as a function of the wave position relative to the source wave are presented in Figure 5.9. There are several features of interest in this velocity profile. First, it can be seen that for short displacements the target wave accelerates when trailing the source wave, as would be expected from the previous examples. For larger displacements, however, there is a reduction in the wave velocity. A simple analysis of this figure allows the determination of the displacement corresponding to the target wave "locked" to the source wave and the range of the initial displacements from which the system converges to this stable configuration. Such locking of the target wave to the moving excitation provided by the source wave is commonly referred to as entrainment.

Figure 5.10 shows the two steady states SS_1 and SS_2 corresponding to the target wave velocity being equal to the source wave velocity. The first steady state SS_1 is unstable, since the small perturbation of the displacement results in the change in the target wave velocity tending to move the wave further in the direction of the initial perturbation. The other steady state SS_2 is stable, with the range of displacements from SS_1 at the left to infinity at the right converging to SS_2 (note $v_2 < v_1$). Thus, the distance between SS_1 and SS_2 determines the stability of the steady state SS_2 with respect to random perturbations on the wave position and as such can serve as a measure stability of the entrainment. In general, this range depends upon both the strength of the coupling between the source and target layers and the difference in their propagation velocities. Figure 5.11 shows the range of stability $|x(SS_1) - x(SS_2)|$
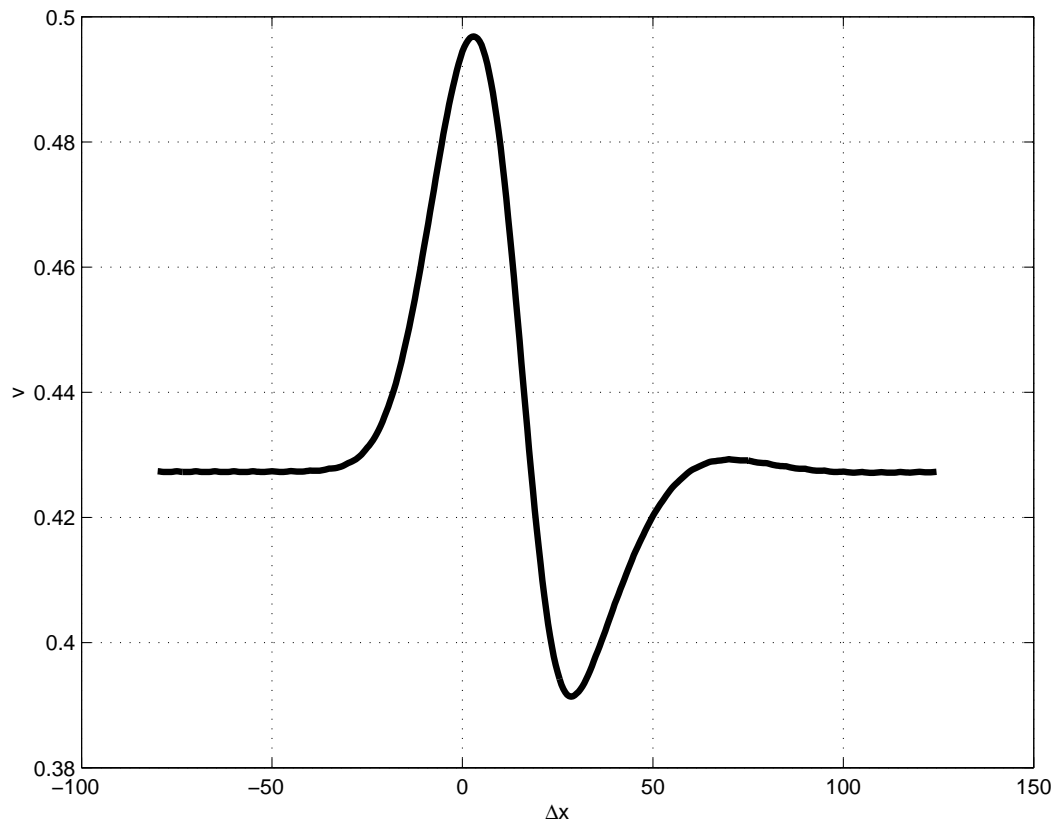
Figure 5.9: The change (dv) in the velocity of the target wave as a function of the relative position (dx) of the target and source waves.
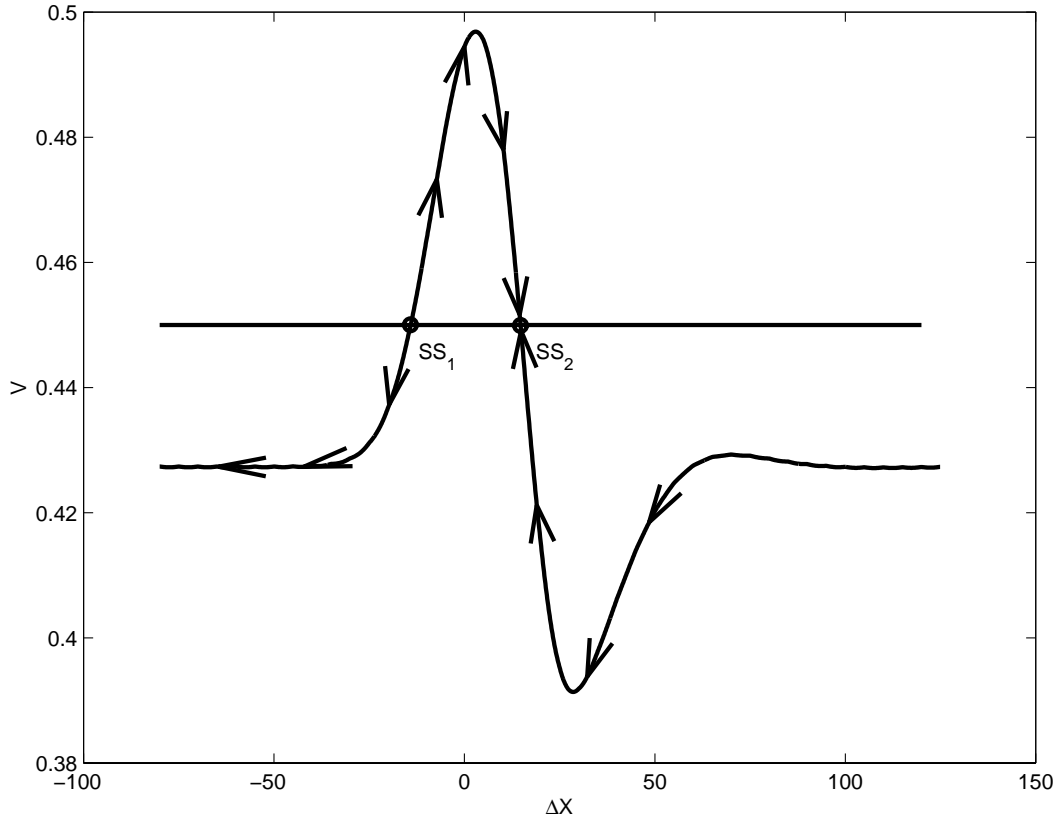
Figure 5.10: The dynamic steady states SS_1 and SS_2 of the wave-to-wave entrainment.

as a function of the mismatch in the intrinsic wave velocity in the layers, as controlled by changing the δ parameter. To a first approximation, the profile of dv does not depend on the difference in the wave velocity and the range can be determined form the Figure 5.10 by drawing the source velocity lines and determining the intersections SS_1 and SS_2 with the dv curve. A similar result is obtained by maintaining δ constant ($\delta = 0.01$) while varying the strength of the coupling between the layers (Figure 5.12). In this case, the source velocity line and the base of the dv curve remain fixed, while the hight of the curve is changing, leading to a similar relationship.

The entrainment is not limited to the case of one-to-one connections. It is not surprising that a slight spreading out of the excitation does not cause qualitative changes in the behavior. Assume that every source neuron connects to a multiplicity of target neurons, with the strength of the connections following gaussian profile G, centered around the source location (Figure 5.13). To a first approximation, the velocity plots
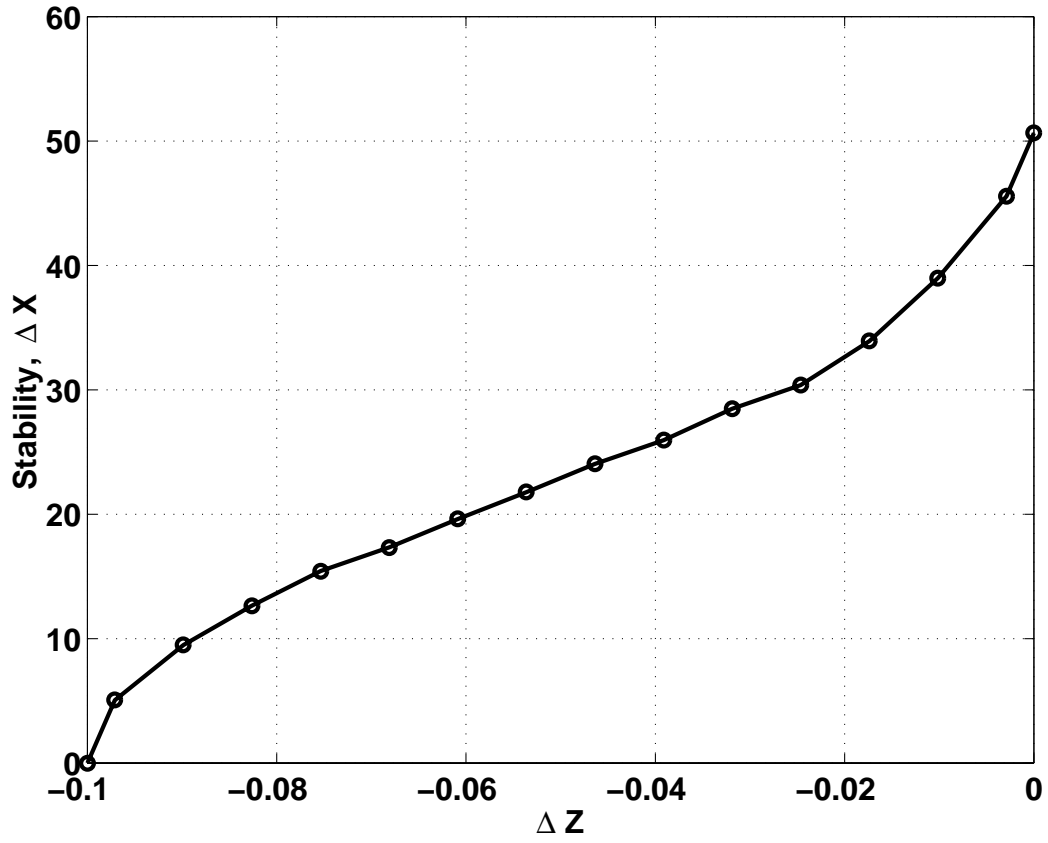
89

Figure 5.11: The region of stable entrainment as a function of ΔZ, the mismatch in the excitatory input to the source and target layers.
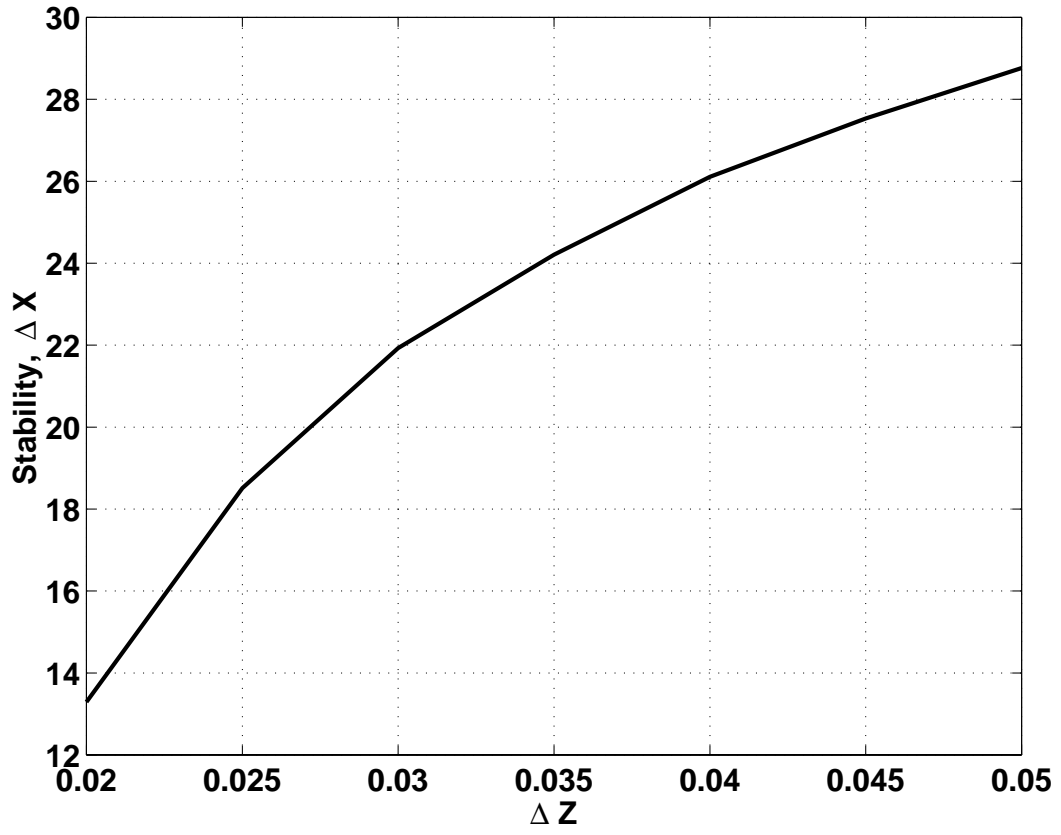
Figure 5.12: The region of stable entrainment as a function of ϵ, the excitatory input to the target layer from the source layer.
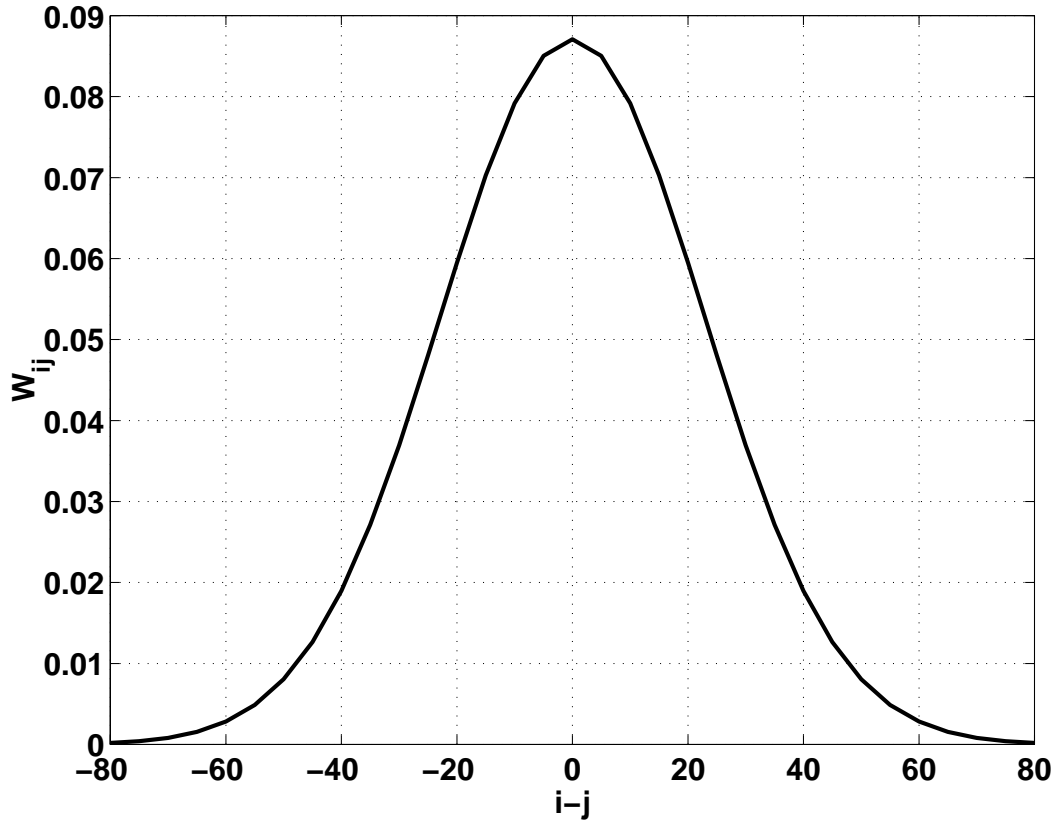
Figure 5.13: The distribution of the weights $\mathbf{W}_{ij}$ as a function of $(i - j)$.

presented above change by a convolution with G. Figure 5.14 demonstrates this effect by comparing the original dv profile convolved with the gaussian G to that obtained by direct measurement of the system with a gaussian-spread excitation. It is worth mentioning that high spatial frequency variations in the strength of the connections between layers have a rather minor effect on the behavior of the system, owing to the low-pass properties of the diffusive coupling between the neighboring neurons.

Comparing the projection of the activity of the input layer to the output layer to the activity elicited in the output layer, the following observation can be made: The projection depends upon the distribution and strength of the connections between the layers, while the activity also involves the influence of the target dynamics. As demonstrated above, in the case of identical or nearly identical dynamics of the layers, there will be dramatically better correspondence between the source and target activity as compared to the source activity and its projection to the target. Based
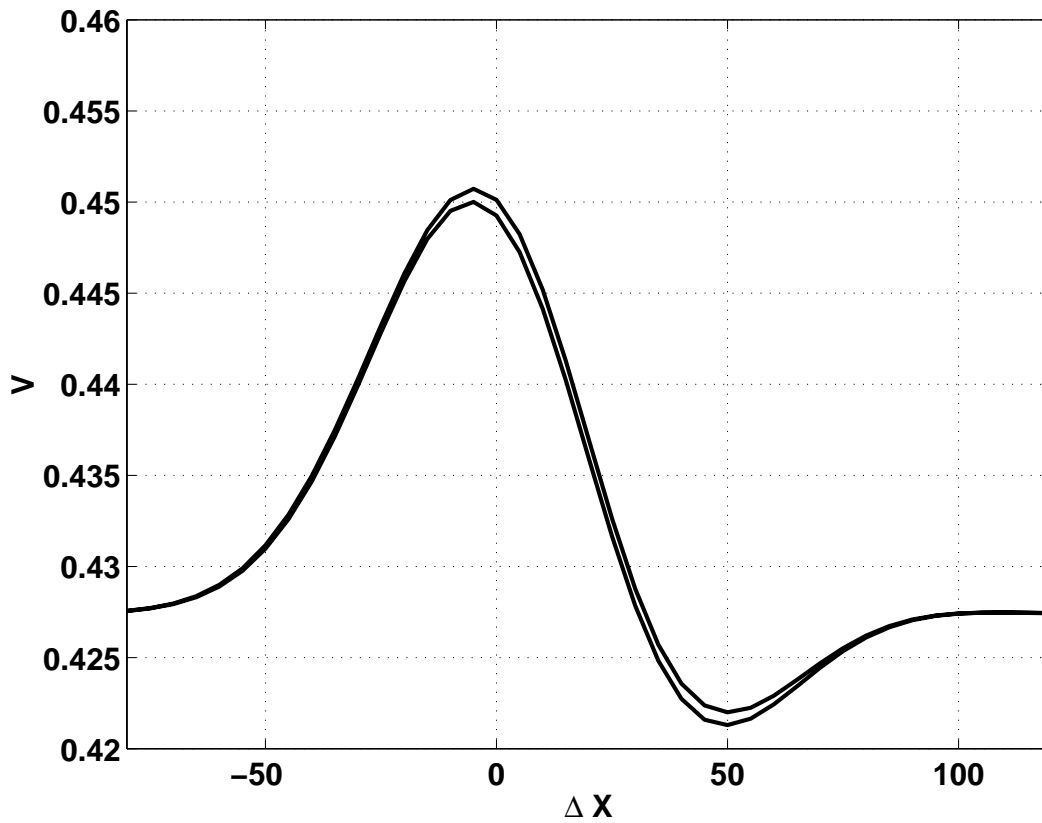
Figure 5.14: Similar to the Figure 5.9 - the overlay of the target velocity profile *dv* for the gaussian-spread excitation (solid line) and the result of the convolution of the original (Figure 5.9) profile with the weight distribution function (Figure 5.13) (dotted line).

on this observation, we suggest that this mechanism can provide matching patterns of activity in two imperfectly connected neural layers. Further we shall demonstrate that these patterns can serve as a reference for the establishment of the topologically precise connections between the layers.

5.3   Model

We consider the development of connections between two layers of neurons. A biological example of such a system is the photoreceptors in the eye that transmit information to the LGN several inches away. The input (**I**) and output (**O**) layers are each modeled as a one-dimensional string ($N = 128$) of Fitzhuh-Nagumo neurons with diffusively coupled nearest neighbors:

$$
\begin{aligned}
\dot{x}_j &= \epsilon \cdot \left(y + x_j - \frac{x_j^3}{3} + z \right) + D_u \cdot \left(x_{j+1} + x_{j-1} - 2x_j \right), \\
\dot{y}_j &= \frac{1}{\epsilon} \cdot \left(a^I - x_j - by_j \right), \\
\dot{u}_j &= \epsilon \cdot \left(v + u_j - \frac{u_j^3}{3} + z \right) + D_u \cdot \left(u_{j+1} + u_{j-1} - 2u_j \right) + \sum_{i=1}^{N} \mathbf{W}_{ij} x_i, \\
\dot{v}_j &= \frac{1}{\epsilon} \cdot \left(a^O - u_j - bv_j \right),
\end{aligned}
\tag{5.4}
$$

with $\epsilon = 3.0$, $\alpha_I = 0.62$ for **I** layer (x, y), $\alpha_O = 0.38$ for **O** layer (u, v), $b = 0.8$, and $D_x = 1.0$; $j = 1..N$. The diffusion term is modified for $j = 1, N$ to provide no-flux boundary conditions. We include unidirectional excitatory connections from each element $\mathbf{I}_i$ to every element $\mathbf{O}_j$ with the strength $\mathbf{W}_{ij}$ (Figure 5.15). It is known that the precision of the initial sketch of the connections, established by the gradients of the chemical markers that are present in both the source and target layers, is orders of magnitude lower than that of the final precision. Thus, we start with randomized initial distribution of $\mathbf{W}_{ij}$, with a controlled bias toward the correct location (ideal $\mathbf{W}_{ij} = \delta_{ij}$), accounting for the marker-based initial layout:

$$
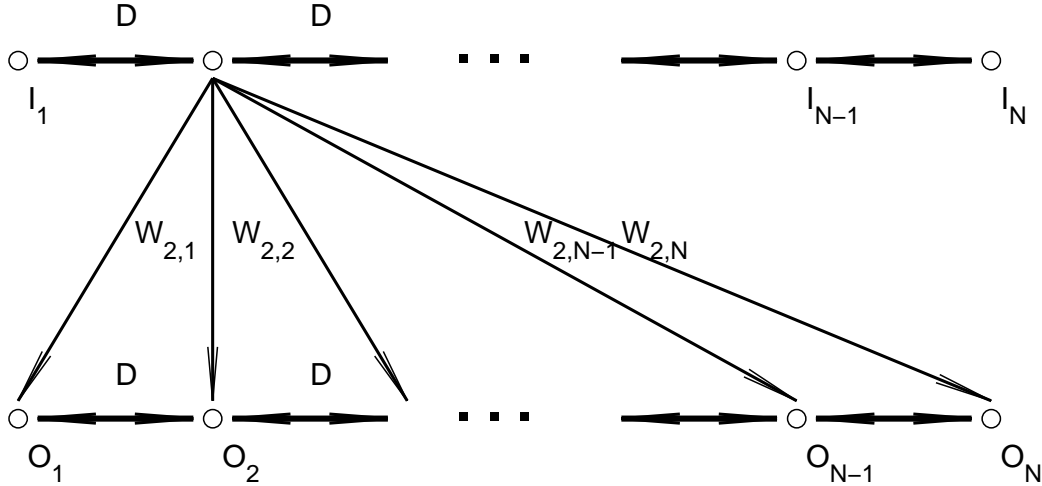\dot{\mathbf{W}}_{ij}^0 = \frac{r}{|i - j + f(j)| + s},
\tag{5.5}
$$

94

Figure 5.15: Two one-dimensional arrays of excitable elements with diffusive coupling between neighbors and excitatory connections from each element $\mathbf{I}_i$ of the source layer to every element $\mathbf{O}_j$ of the target layer with the strength $\mathbf{W}_{ij}$.

where r are random numbers drawn from a uniform distribution on the interval $0..1$, $s = 6.0$ is a parameter controlling the half-width of the distribution, and $f(j) = 10.0\sin(2\pi j/N)$ is arbitrarily chosen function to simulate distortion in the initial layout of the connections. [For a representative $\mathbf{W}_0$, see Figure 5.16]

It is known that the unscrambling of initially imperfect connections in the optic nerve requires visual input. Pre-birth spontaneous activity, observed experimentally in the retina, has been suggested as the necessary input early in development.[12, 19] To simulate this spontaneous wave activity, layer $\mathbf{I}$ is infrequently stimulated (at 40s intervals) at random locations to initiate individual waves of activity. Layer $\mathbf{O}$ will occasionally become activated at certain locations due to the excitation received through the connections. The initiated waves will propagate across this layer in a manner similar to that in the layer $\mathbf{I}$. The simultaneous excitation of every pair $\mathbf{I}_i$, $\mathbf{O}_j$ causes a small increase in $\mathbf{W}_{ij}$, accounting to the Hebbian rule:

$$\dot{\mathbf{W}}_{ij} = 0.01 \ \textit{if} \ u_i < 0 \ \& \ u_j < 0. \tag{5.6}$$

The total sum of all of the weights converging to a given cell is kept constant by periodic normalization in order to account for "unlearning" and other negative feedback processes not explicitly incorporated in the model:[18]

95
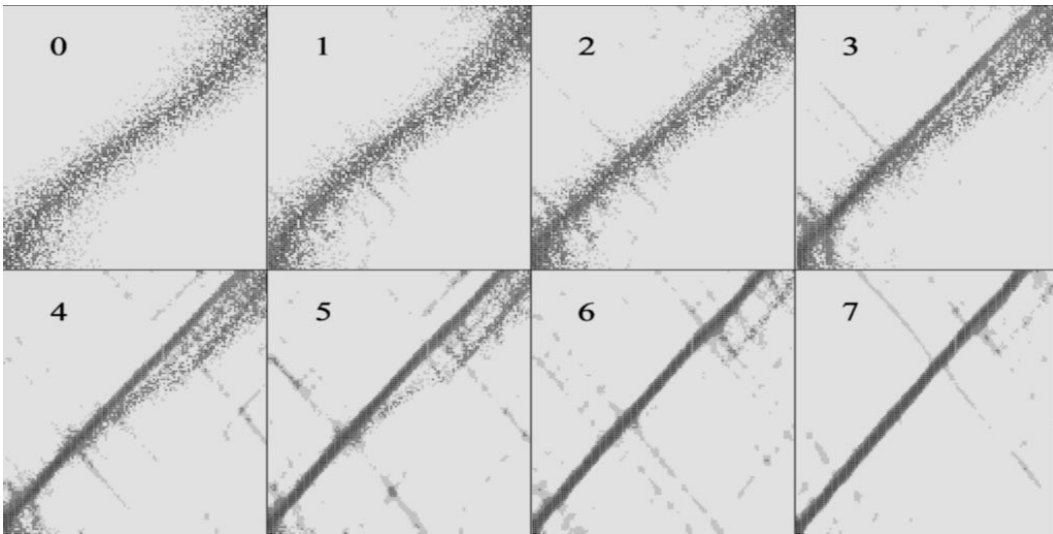
Figure 5.16: The weight matrix **W** is represented as a grayscale image with intensity at each (x, y) location corresponding to the magnitude of a $\mathbf{W}_{yx}$; Successive images are taken every 100 epoch of simulation starting from the initial distribution.

Figure 5.17: A grayscale image is applied as the input to the layer **I** and excitation projected using the current weight matrix **W** to the output layer **O** is shown. Successive output images correspond to the progressively evolving **W** weight matrix as presented above on Figure 5.16.

$$\sum_{i=1}^{N} \mathbf{W}_{ij} = 1. \tag{5.7}$$

Every 1000 seconds of simulated time, the weight matrix is evaluated by applying a known pattern to the **I** layer and computing the excitation projected onto the **O** layer. Individual lines of a photographic image were used as input patterns, and the output was reassembled into a two-dimensional array, allowing visual inspection of the extent of scrambling occuring in the transmission from the **I** to **O** layers.

5.4 Results

The local-scale disorder in the connections makes the activation reaching the output different from the pattern of activity in the input layer. However, on a coarser scale there is correlation between ongoing and projected activities, because coarse-scale topology can be established by developmental mechanisms. Under such conditions, activity elicited in the output layer, on average, resembles ongoing activity in the input layer much closer than would be expected from the pattern of the received activation. The two layers synchronize their patterns of activity under relatively weak, diffuse coupling. This synchronization allows the Hebbian mechanism to adjust the existing scheme of connections to a higher degree of topological correspondence. The effect of this adjustment is an increased correlation between the activation reaching the output layer and the pattern of activity in the input layer. This, in turn, allows for an even higher degree of synchronization between the patterns of activity in the input and output layers, and so on, until a nearly perfect match is established.

The analysis of the behavior of this model provides a new perspective on the possible mechanisms for the establishment of a precise topology of connections in the visual pathways and other neural tracts between isomorphic areas. In particular, the model provides insights into the mechanism responsible for the organization of proper connections in the "receiving" layer. We draw two conclusions from our results: (1) Entrainment of the receiving circuit by the source circuit, on the basis of initially weak and diffuse coupling, can provide matching patterns of activity in both circuits, allowing plasticity mechanisms to establish more precise connections, and (2),

entrainment and plasticity can behave in a closed - loop manner, resulting in fast, iterative improvement of an initially imprecise scheme of connections. Weak, diffusive coupling with Hebbian learning offers an attractive mechanizm for the development of neurophysiological connections.

5.5   Conclusion

A number of experimental studies have demonstrated the importance of spontaneous activity in different areas of visual,[19, 20] auditory,[8] and other parts of neural system[9] for the development of appropriate neural connections. It has been clearly demonstrated that genetically encoded labels provide guidance for the diffuse connections, but no more than a partial bias towards the correct locations,[22–24] while the onset of synaptic activity leads to a rapid reorganization of the connections, resulting in precise and topologically correct projections.[25] Similarly, in the mature nervous system, spontaneous activity during certain sleep phases has been implicated in maintaining the pattern of existing connections, thus stabilizing the structure with respect to adverse environmental inputs. While it is certain that the details of spontaneous, pattern-driven self-organization vary in different systems due to anatomical differences, there is a striking similarity between all these phenomena and certain aspects of the behavior of coupled spatially extended chemical systems. Entrainment observed experimentally in a chemical system containing two active layers[26] offers a possible mechanism for matching patterns of activity in source and target systems, as well as serving as a guide in establishing and maintaining topographic projections between the connected layers. In this chapter we have initiated an exploration of the interactions between entrainment and Hebbian mechanisms. The model presented here, which can be viewed as a caricature of a neural system, such as the eye – LGN circuit, exhibits rapid convergence to a topological order in a system of initially random connections, without the use of any specialized mechanisms. Such behavior, although not directly comparable to any biological system due to the crudeness of the model, offers a possible explanation of the observed developmental processes on the gross level. Inclusion of the structural details and use of a more realistic neuron model may enable more direct comparisons with real neural circuits (at the expense, however,

of conceptual clarity). Another possible extension of the model lies in the possibility of synchronizing chaotic systems[27] without requiring detailed topological correspondence. In this case, the term "projection" would acquire additional meaning, since such systems would perform projections from the state space of the source circuit to the state space of the target circuit, thus performing information processing equivalent to the nonlinear transformation from one coordinate system to another. The relevance of such circuits to neural processing is not clear at this time; however, such circuits seem to be biologically possible, and it also should be noted that an extension of the "filtering" technique presented in the previous chapter, using mesh-based interpolation as the basis for an Extended Kalman Filter, could be implemented as such a circuit. Exploration of this possibility may provide an explanation of the mechanisms responsible for the separation of signals, such as source separation in the "cocktail party" situation and figure-background separation in the case of visual inputs, which are among the fundamental elements of the recognition processes.

5.6 Appendix A. Brain is not directly encoded

The human genome contains $n \approx 3 \cdot 10^9$ nucleotide pairs.[28] Since each of these locations contains one out of the four possible nucleotides, it carries $m = 2$ bits of information ($\log_2(4)$), giving the complete genome the information capacity $mn = 6 \cdot 10^9$ bits. It is believed that 95% to 99% of the chromosomal DNA serves structural roles and does not encode any genes. Furthermore, the coding part is substantially redundant, frequently containing multiple copies of the same gene. These facts place an upper bound on the order of 10^8 bits to the amount of useful genetic information contained in the humane genome.

It is estimated that the number of neurons in the brain reaches $n \approx 10^{11}$, with each neuron making from a few hundred to a few hundred thousand connections with an average of about $m \approx 10^4$.[4] The amount of information necessary to describe the connectivity of this structure is at least $nm\log_2(n) > 10^{16}$. This number does not take into account the variable strength of the connections, the representation of which would require a great deal of additional information. It also neglects the additional complexity that connections to a specific neuron can occur at various parts of its

soma and dendritic tree, also requiring specification. These factors may increase the above estimate by several orders of magnitude. Comparing the numbers estimated from the genetic information and the complexity of the nervous system, it is obvious that the genetic material cannot possibly specify each individual neural connection independently, and thus, the structure of the brain cannot be directly encoded in the genome.

The evidence from experiments on flies and cockroaches demonstrates that structure of the neural systems in these organisms is highly conservative.[5] The neural structure is almost exactly the same in individual insects with identical genetic makeup. The genetic material of a fly is ≈ 0.1 of the human[28] and its nervous system contains about 10^3 times less neurons, which, in analogy with the above estimates, corresponds to 10^7 bits and more than 10^{13} bits, respectively. The mismatch between the amount of genetic material and the complexity of the neural structure is similar to that in the case of the human species. However, we are certain that the structure of the nervous system is very precisely defined by the genetic material. Thus the structure of the nervous system must be presented in the genetic material in an efficiently compressed form.

5.7   Appendix B. Numerical Method

The simulations were performed using the following scheme:

$$x_{n+1} = x_n + hf(x_{n+1/2}). \tag{5.8}$$

To find the value of the derivative at the interval midpoint we proceed as follows:

$$\begin{aligned} x_{n+1/2} &= x_n + hf(x_{n+1/2}), \\ x_{n+1} &= x_{n+1/2} + hf(x_{n+1/2}). \end{aligned} \tag{5.9}$$

The solution of the nonlinear equation arising from the implicit step is found by Newton-Raphson iterations:[29]

$$x_n^{j+1} = x_{n-1}^{\infty} + h\{f(x_n^j) + f(x_n^j)(x_n^{j+1} - x_n^j)\}, \tag{5.10}$$

thus

$$\Delta x_n = [E - hJ(x_n)]^{-1} \{x_{n-1} + hf(x_n) - x_n\}, \tag{5.11}$$

where

$$[E - hJ]^{-1} = \frac{1}{D} \begin{bmatrix} 1 + bh/c & -h/c \\ -h/c & 1 + ch(x^2 - 1), \end{bmatrix} \tag{5.12}$$

and

$$D = \{1 + ch(x^2 - 1)\}\{1 + bh/c\} + h^2. \tag{5.13}$$

5.8 References

[1] W. J. Gehring, "The Master Control Gene for Morphogenesis and Evolution of the Eye," *Genes to Cells* **1**, 11-15, (1996).

[2] P. P. Tam and R. R. Behringer, "Mouse Gastrulation: the Formation of a Mammalian Body Plan," *Mechanisms of Development* **68**, 3-25 (1997).

[3] T. Poggio and W. Reichardt, "Characterization of Non-linear Interaction in the Fly's Visual System," in T. Poggio and W. Reichardt, eds., *Theoretical Approaches in Neurobiology,* (M.I.T. Press, Cambridge, MA, 1981); A. Borst, "How do Nerve Cells Compute? Dendritic Integration in Fly Visual Interneurones," *Acta Physiologica Scandinavica* **157**, 403-407, (1996).

[4] E. R. Kandel, J. H. Schwartz, and T. M. Jessell, eds., *Principles of Neural Science,* (Appleton and Lange, 1991).

[5] C.A.G.Wiersma, ed., *Invertebrate Nervous Systems,* (University of Chicago Press, 1967).

[6] S. I Amari, "Topographic Organization of Nerve Fields," *Bulletin of Mathematical Biology* **42** 339-364 (1980).

[7] R. J. Weinberg, "Are topographic maps fundamental to sensory processing ?" *Brain Research Bulletin* **44**, 113-116 (1997).

[8] R. Romand, "Modification of Tonotopic Representation in the Auditory System During Development," *Progress in Neurobiology* **51**, 1-17 (1997).

[9] M. J. O'Donovan, N. Chub, and P. Wenner, "Mechanisms of Spontaneous Activity in Developing Spinal Networks," *Journal of Neurobiology* **37** 131-145 (1998).

[10] S. B. Udin, "Development of Orderly Connections in the Retinatectal System," in K. N. Leibovic, ed., *Science of Vision,* (Springer-Verlag NY, 1990).

[11] I. Thompson, "Visual Development. From Darkness into Light," *Curr. Biol.* **4**, 458-461 (1994).

[12] M. B. Feller, D. P. Wellis, D. Stellwagen, F. S. Werblin and C.J. Shatz, "Requirement for Cholinergic Synaptic Transmission in the Propagation of Spontaneous Retinal Waves," *Science* **272**, 1182-1187 (1996).

[13] D. O. Hebb, *The Organization of Behavior,* (John Wiley, New York 1949).

[14] A. L. Hodgkin and A. F. Huxley, "A Quantitative Description of Membrane Current and its Application to Conduction and Excitation in Nerve," *J. Physiol. London*, **117**, 500-544, (1952).

[15] D. Jonhston and S. Miao-Sin Wu., *Foundations of Cellular Neurophysiology,* (MIT Press, Cambridge, Massachusetts, 1995).

[16] R. A. FitzHugh, "Impulses and Physiological States in Theoretical Models of Nerve Membrane ," *Biophysical Journal* **1**, 445-466 (1961).

[17] G. D. Smith, *Numerical Solution of Partial Differential Equations : Finite Difference Methods,* (Oxford Univ Press, 1986).

[18] G. J. Goodhill and H. G. Barrow, " The Role of Weight Normalization in Competitive Learning," in *Neural Computation* **6**, 255-269 (MIT Press 1994).

[19] L. C. Katz and C. J. Shatz, "Synaptic Activity and the Construction of Cortical Circuits," *Science* **274**, 1133-1138 (1996).

[20] A. Penn, P. Riquelme, M. Feller, C. Shatz, "Competition in Retinogeniculate Patterning Driven by Spontaneous Activity," *Science* **279**, 2108-2112 (1998).

[21] C. E. Holt, W. A. Harris, "Position, Guidance, and Mapping in the Developing Visual System," *J. Neurobiol.* **24**, 1400-1422 (1993).

[22] D. A. Feldheim et al., "Topographic Guidance Labels in a Sensory Projection to the Forebrain," *Neuron* **21**, 1303-1313 (1998).

[23] J. Yuasa, S. Hirano, M. Yamagata, and M. Noda, "Visual projection map Specified by Topographic Expression of Transcription Factors in the Retina," *Nature* **382**, 632-635 (1996).

[24] D. K. Simon, D. D. O'Leary, "Limited Topographic Specificity in the Targeting and Branching of Mammalian Retinal Axons," *Dev. Biol.* **137**, 125-134 (1990).

[25] J. Alonso, W. Usrey, and R. Reid, "Precisely Correlated Firing in Cells of the Lateral Geniculate Nucleus," *Nature* **383**, 815-819 (1996).

[26] D. Winston, M. Arora, J. Maselko, V. Gaspar and K. Showalter, "Cross-Membrane Coupling of Chemical Spatiotemporal Patterns," *Nature* **351**, 132-135 (1991).

[27] L. Pecora and T. Caroll, "Synchronization in Chaotic Systems," *Phys. Rev. Lett.* **64**, 821-823 (1990).

[28] B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, J. D. Watson, *Molecular Biology of the Cell,* (Garland Pub, 1995).

[29] W. Press, B. Flannery, S. Teukolsky, and W. Vetterling, *Numerical Recipes: The Art of Scientific Computing,* (Cambridge University Press: Cambridge, 1986).