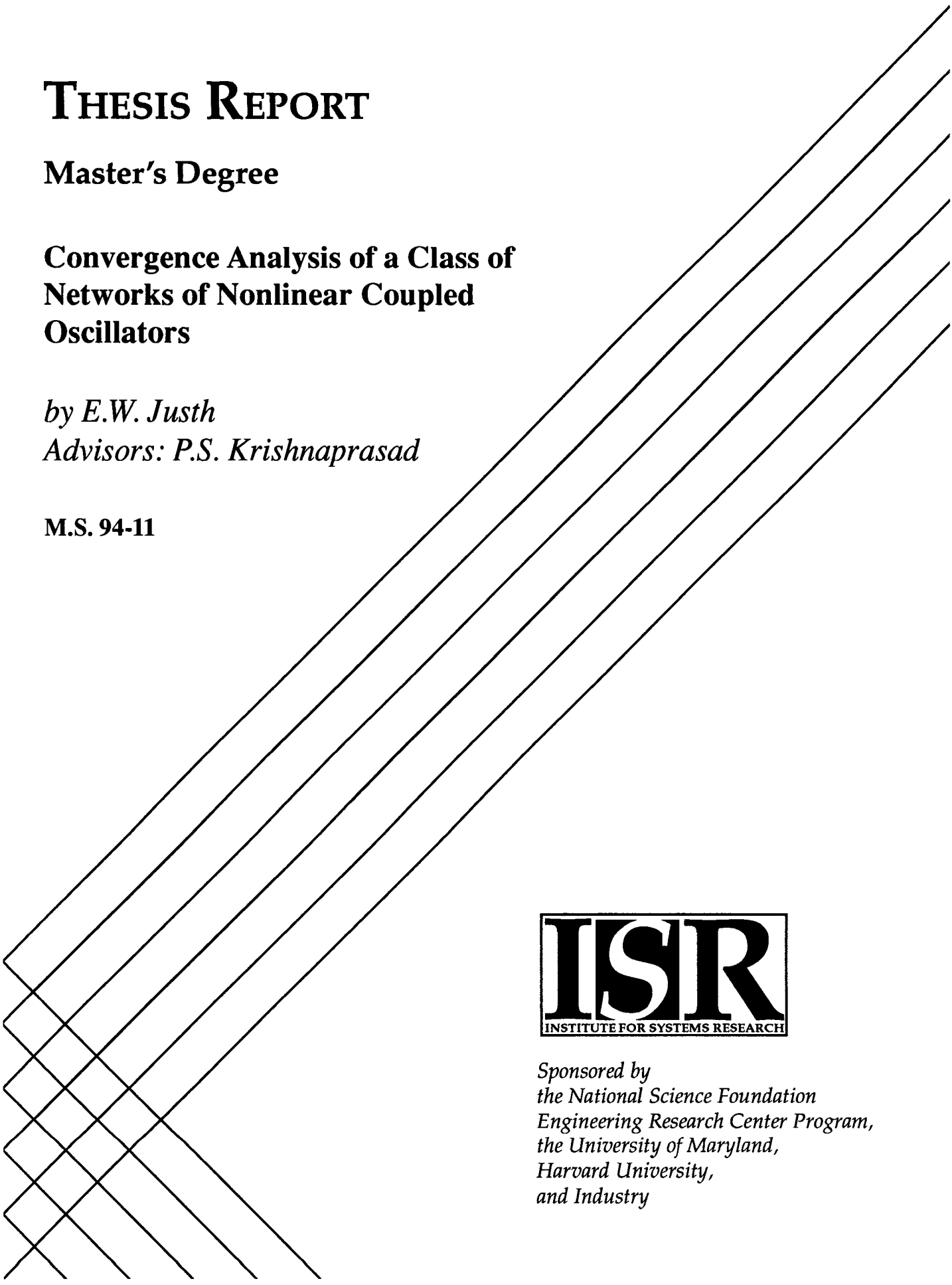
THESIS REPORT

Master's Degree

Convergence Analysis of a Class of Networks of Nonlinear Coupled Oscillators

by E.W. Justh
Advisors: P.S. Krishnaprasad

M.S. 94-11



Sponsored by
the National Science Foundation
Engineering Research Center Program,
the University of Maryland,
Harvard University,
and Industry

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
1994		**00-00-1994 to 00-00-1994**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Convergence Analysis of a Class of Networks of Nonlinear Coupled Oscillators	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of Maryland,The Graduate School,2123 Lee Building,College Park,MD,20742**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
see report

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**97**	

ABSTRACT

Title of Thesis: CONVERGENCE ANALYSIS OF A CLASS OF NETWORKS OF NONLINEAR COUPLED OSCILLATORS

Name of degree candidate: Eric W. Justh

Degree and Year: Master of Science, 1994

Thesis directed by: Professor P.S. Krishnaprasad
Department of Electrical Engineering

A network of nonlinear coupled oscillators is presented, and a convergence proof is given along with physical motivation. Next, the network architecture is generalized by allowing interconnections between oscillators to be controlled in an adaptive fashion, and convergence of the generalized network is proved. An example network is presented to illustrate the utility of such networks and to show why the problem of undesired stable equilibria must be addressed. Two alternative approaches are then presented which overcome the problem of undesired stable equilibria appearing in the network dynamics. Finally, an analog VLSI approach to implementation of such networks is presented, and tradeoffs among power dissipation, bandwidth, and network size are discussed.

CONVERGENCE ANALYSIS OF A CLASS OF NETWORKS OF NONLINEAR COUPLED OSCILLATORS

by
Eric W. Justh

Thesis submitted to the Faculty of the Graduate School
of The University of Maryland in partial fulfillment
of the requirements for the degree of
Master of Science
1994

Advisory Committee:
Professor P.S. Krishnaprasad, Chairman/Advisor
Professor W.P. Dayawansa
Professor S. Shamma

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. P.S. Krishnaprasad, for his help and guidance on the theoretical aspects of this thesis. I would also like to thank Dr. Fritz Kub, Head of the Microelectronic Device Physics Section at the U.S. Naval Reseach Laboratory. Under Dr. Kub, I worked on the analog VLSI high-speed adaptive filter project, which provided insight into the implementation aspects of this thesis.

This research was supported in part by the National Science Foundation's Engineering Research Centers Program: NSFD CDR 8803012, and by the AFOSR University Research Initiative Program, under grant AFOSR-90-0105 and AFOSR-F49620-92-J-0500. Also, this work was partially supported by a National Science Foundation Graduate Fellowship.

TABLE OF CONTENTS

LIST OF FIGURES

NUMBER PAGE

Chapter 1

Introduction

1.1 Background

Networks of coupled oscillators have been proposed for various applications, including locomotion (specifically, central pattern generators), pattern recognition (for instance, distinguishing between different objects in view simultaneously), and biological information processing (as a solution to the dynamic binding problem). There is evidence that coupled oscillator networks play a role in biological systems, but the point of view taken in this work is that the mathematical analysis of coupled oscillator networks can also yield networks which are useful in their own right for applications in pattern recognition, control, and communications.

Experiments on the spinal cord of the lamprey [1], and the attempts to model the experimental results using networks of coupled oscillators [2][3][4] are a good illustration of how relatively simple coupled oscillator networks can model a variety of observed behaviors in an actual biological motor-control network. There has also been recent work in designing control systems based

on coupled oscillators for both hexapod and biped robotic locomotion [5][6].

Experiments on cats have revealed synchronous firings of neurons in the cat visual cortex in response to visual stimulus [7][8]. Moreover, there has been work in developing artificial pattern recognition systems based on networks of coupled oscillators, in which synchronous oscillations serve as a dynamic binding mechanism [9][10][11]. Moreover, a method of performing systematic reasoning based on using synchronous oscillations as a dynamic binding mechanism has been proposed, both for artificial information processing and as a model sharing some features of human reflexive reasoning [12].

Besides biologically inspired uses for coupled oscillator networks, there are control problems for which periodic inputs have been proposed and studied [13] [14]. Central pattern generators based on networks of coupled oscillators may prove useful as control systems for such problems.

1.2 Overview

The basic network examined in this work is given by

$$\dot{y}_j = -y_j + r(\beta|x_j|)\frac{x_j}{|x_j|} \tag{1.1}$$

$$x_j = \sum_{k=1}^{n} y_k w_{jk}^*, \quad j = 1, ..., n \tag{1.2}$$

where x_j and y_j are complex numbers $\forall j$, w_{jk} are complex interconnecting weights with $w_{jj} = 0 \ \ \forall j$ and $w_{jk} = w_{kj}^*$ (i.e., the weight matrix is Hermitian), β is a scalar parameter, and $r(\cdot) : [0, \infty) \rightarrow \Re$ is a memoryless strictly monotone increasing nonlinearity with $r(0) = 0$ and $\lim_{m\rightarrow\infty} r(m) = 1$ (later to be specified precisely). $r(\cdot)$ is also assumed to be analytic, and the function $r(\beta|z|)\frac{z}{|z|} : \mathbb{C} \rightarrow \mathbb{C}$ is a well-defined function which compresses the magnitude of its complex argument while retaining its angle.

For the j^{th} unit (or oscillator), y_j may be thought of as its state, and x_j represents its input from the rest of the network. Because the states are complex, each unit carries both phase and amplitude information, and it is the phase information which is of primary interest. In the coupled oscillator context, the relative phases of a network of oscillators oscillating at the same frequency are represented.

The reason for choosing these dynamical equations is that they repre-

sent a very simple method of defining a coupled-oscillator network. The input to each unit is the weighted sum of the states of the other units, where the weights themselves are complex (and thus may alter the phase as well as the amplitude of the signals they weight). The input to each unit is then passed through a memoryless saturating nonlinearity whose gain is a parameter, and the result is then low-pass-filtered to produce that unit's state.

In Chapter 2, convergence of this network, in the spirit of LaSalle's Invariance Principle, is proved, and physical motivation for the proof is given. Next, in Chapter 3, the weights (w_{jk}) are allowed to be units (y_l) themselves, and convergence of an adaptive control law for such systems is proved. In Chapter 4, a 6-unit example circuit is presented illustrating how undesired stable equilibria can arise. Then two alternative approaches are then presented which overcome the problem of undesired stable equilibria appearing in the network dynamics. Implementation issues are addressed throughout, and finally, an analog VLSI approach for implementing such networks is presented in Chapter 5. In Chapter 6 actual VLSI circuits which could be used to build a coupled oscillator network are described in order to permit the tradeoffs among power dissipation, bandwidth, and network size to be explored. Chapter 7 concludes with recommendations for future work.

Chapter 2

Fixed-Weight Network Analysis

2.1 Introduction

The dynamics under consideration are:

$$\dot{y}_j = -y_j + r(\beta|x_j|)\frac{x_j}{|x_j|}, \quad y_j \in \mathbb{C} \tag{2.1}$$

$$x_j = \sum_{k=1}^{n} y_k w_{jk}^*, \quad x_j \in \mathbb{C}, \quad j = 1, ..., n \tag{2.2}$$

where $w_{jk} = w_{kj}^* \in \mathbb{C}$, $w_{jj} = 0 \ \ \forall j$, $\beta \in \Re$ is constant, and $r(\cdot) = \frac{I_1(\cdot)}{I_0(\cdot)}$. Here, I_0 and I_1 are the modified Bessel functions of the first kind and of order zero and one, respectively, and $r(\cdot) : [0, \infty) \to \Re$, is a memoryless strictly increasing saturating nonlinearity with $r(0) = 0$ and $\lim_{m\to\infty} r(m) = 1$. (The choice of this particular saturating nonlinearity will be motivated later).

The first step in the convergence analysis is to rewrite the dynamics in terms of the x_j variables:

$$\begin{aligned}
\dot{x}_j &= \sum_k \dot{y}_k w_{jk}^* \\
&= \sum_k \left(-y_k + r(\beta|x_k|)\frac{x_k}{|x_k|} \right) w_{jk}^* \\
&= -\sum_k y_k w_{jk}^* + \sum_k r(\beta|x_k|)\frac{x_k}{|x_k|} w_{jk}^*
\end{aligned}$$

$$= -x_j + \sum_k r(\beta|x_k|)\frac{x_k}{|x_k|}w^*_{jk}. \tag{2.3}$$

These dynamics were proposed and physically motivated by Zemel, Williams, and Mozer [15]; however, their proof of convergence was incomplete. After a review of the physical motivation, a rigorous proof of convergence of their network using LaSalle's Invariance Principle will be presented.

2.2 Physical Motivation

Zemel, Williams, and Mozer propose first a stochastic network of directional units (complex-valued random variables with magnitude one and angle representing directional information) interconnected by complex weights which are considered fixed [15]. The directional units evolve according to probability distributions determined by the other directional units and interconnecting weights. The stochastic network is then simplified using the mean-field approximation to give a deterministic network. The purpose of examining the stochastic network is that it provides insight for the stability analysis of the deterministic network.

To begin the stochastic network analysis, consider a network of directional units, each represented by a random variable Z_j taking values on the unit circle in the complex plane. The directional units are interconnected

by fixed complex weights w_{jk} satisfying $w_{jj} = 0$ and $w_{kj} = w_{jk}^*$. Defining $x_j = \sum_k z_k w_{jk}^*$ to represent the interaction of unit j with the rest of the network, the mean value of Z_j is determined by the angle of x_j, and the variance of Z_j is inversely related to the magnitude of x_j (the precise dependence to be determined below).

Next, a quadratic form representing "energy" in terms of unit states and interconnecting weights is defined:

$$E(\mathbf{z}) = -\frac{1}{2}\mathbf{z}^T\mathbf{W}\mathbf{z}^* = -\frac{1}{2}\sum_{j,k} z_j z_k^* w_{jk}, \quad \mathbf{z} = (z_1, ..., z_n)^T. \tag{2.4}$$

Because $\mathbf{W}$ is hermitian (i.e., $w_{kj} = w_{jk}^*$), $E(\mathbf{z})$ is real-valued. (This definition of energy generalizes the Hopfield energy function for binary units [15].) Using $x_j = \sum_k z_k w_{jk}^*$ and changing to polar coordinates: $x_j = a_j e^{\imath\alpha_j}$ and $z_j = e^{i\phi_j}$ (recall that $|z_j|$ is constrained to equal 1), we define

$$\begin{aligned} E_j(\mathbf{z}) &= -\frac{1}{2}\left[z_j x_j^* + (z_j x_j^*)^*\right] \\ &= -a_j \cos(\phi_j - \alpha_j) \end{aligned} \tag{2.5}$$

as unit j's contribution to the total energy. Then $E(\mathbf{z}) = \frac{1}{2}\sum_j E_j(\mathbf{z})$. (Observe that when the angle of x_j is aligned with the angle of z_j, unit j is in a low-energy state.) Introducing a "Boltzmann factor" β (interpreted as the reciprocal of temperature), and taking the probability that the jth unit is in

a state $z_j = e^{i\phi_j}$ to be proportional to $e^{-\beta E_j(\mathbf{z})}$, we obtain:

$$f_{Z_j}(z_j) \propto e^{\beta a_j \cos(\phi_j - \alpha_j)} \tag{2.6}$$

where Z_j, the state of unit j, is a random variable taking values on the unit circle in the complex plane and $f_{Z_j}(\cdot)$ is a probability density function. Appropriate normalization gives

$$f_{Z_j}(\phi_j) = \frac{1}{2\pi I_0(m_j)} e^{m_j \cos(\phi_j - \bar{\phi}_j)}, \quad m_j = \beta a_j, \quad \bar{\phi}_j = \alpha_j \tag{2.7}$$

where $I_0(\cdot)$ is the modified Bessel function of the first kind and order zero. This is known as the Von Mises, or circular normal, distribution, and it is a distribution for circular random variables having some characteristics similar to the usual normal distribution for linear random variables [16]. A circular normal distribution is completely characterized by two parameters: a mean direction $\bar{\phi} \in [0, 2\pi)$ and a concentration parameter $m > 0$ which corresponds to the reciprocal of the variance of a linear normal random variable.

Next, Zemel, Williams, and Mozer apply a mean-field approximation to come up with a deterministic network model. In the mean-field approximation, the random variables Z_j are replaced by their means $y_j =< Z_j >$ and are treated as independent (even though they are, in fact, highly coupled). The mean $< Z_j >$ of a Von Mises random variable is a complex number $y_j = r_j e^{i\gamma_j}$ with $\gamma_j = \bar{\phi}_j$ and $r_j = \frac{I_1(m_j)}{I_0(m_j)}$. Figure 2.1 shows r_j as a function of

m_j: it is strictly monotone increasing, passes through the origin, and satisfies $\lim_{m_j \to \infty} r_j(m_j) = 1$.

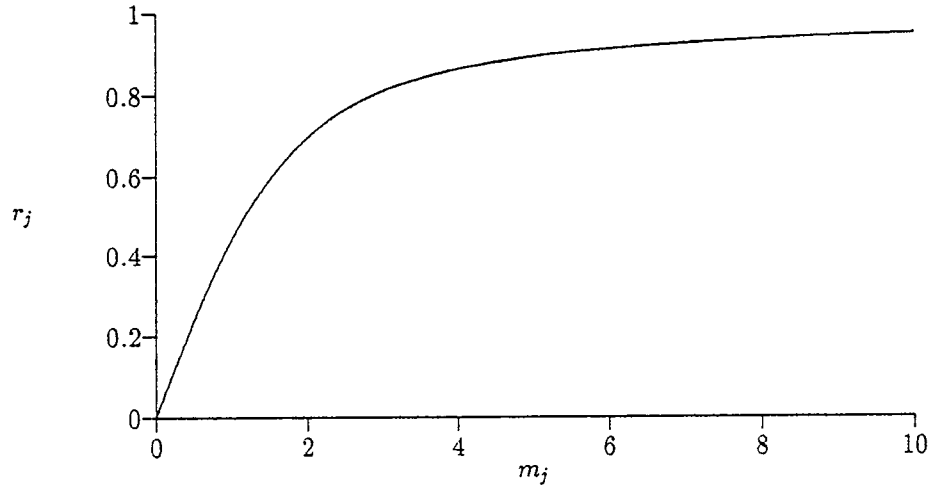


Figure 2.1: Plot of $r_j = \frac{I_1(m_j)}{I_0(m_j)}$ (from [15]).

The dynamics for the deterministic network are chosen to be

$$\frac{dx_j}{dt} = -x_j + \sum_k y_k w^*_{jk} \tag{2.8}$$

so that at equilibrium, $x_j = \sum_k y_k w^*_{jk}$, in analogy with the expression $x_j = \sum_k z_k w^*_{jk}$ for the stochastic network. Furthermore, $y_j = r(\beta|x_j|)\frac{x_j}{|x_j|}$, so that if x_j is viewed as determining the mean and concentration parameter of a Von Mises distribution according to (2.7), y_j will be the mean value of the distribution. The deterministic dynamics can be expressed in terms of the

x_j alone as:

$$\frac{dx_j}{dt} = -x_j + \sum_k r(\beta|x_k|)\frac{x_k}{|x_k|}w^*_{jk}. \tag{2.9}$$

The total energy for the deterministic network is found by taking the mean of the total energy expression for the stochastic network (and assuming the units are independent):

$$< E >= -\frac{1}{2}\sum_{j,k} y_j y^*_k w_{jk}. \tag{2.10}$$

Furthermore, the "entropy" for the deterministic network is found by summing the entropies of the individual units of the stochastic network (again using the independence assumption):

$$H = \sum_j \left[-\beta a_j \frac{I_1(\beta a_j)}{I_0(\beta a_j)} + \log(2\pi I_0(\beta a_j))\right]. \tag{2.11}$$

With these definitions of $< E >$ and H, a Lyapunov function corresponding to what Zemel et. al. call "free energy," $F =< E > -TH$, $T = \frac{1}{\beta}$, can be computed, and this Lyapunov function can be used to prove convergence of the deterministic network using LaSalle's Invariance Principle. The rigorous arguments behind this are a main contribution of this thesis.

2.3 Proof of Convergence

As will now be shown, the deterministic dynamics (2.9) are conver-

gent: every trajectory converges to an equilibrium point. LaSalle's Invariance Principle is invoked to prove this, and a Lyapunov function based on the physically motivated "free energy" is used.

The complex nonlinear dynamics are:

$$\dot{x}_j = -x_j + \sum_k r(\beta|x_k|)\frac{x_k}{|x_k|}w_{jk}^*, \quad j = 1, ..., n \tag{2.12}$$

where the x_j are complex, $r(m) = \frac{I_1(m)}{I_0(m)}$, $w_{kj} = w_{jk}^*$, and $w_{jj} = 0 \ \forall j$. Letting $x_j^R = Re(x_j)$ and $x_j^I = Im(x_j)$ we can rewrite the dynamics as

$$\begin{aligned} \dot{x}_j^R &= -x_j^R + \sum_k r(\beta|x_k|)\frac{x_k^R w_{jk}^R + x_k^I w_{jk}^I}{|x_k|} \\ \dot{x}_j^I &= -x_j^I + \sum_k r(\beta|x_k|)\frac{x_k^I w_{jk}^R - x_k^R w_{jk}^I}{|x_k|}, \quad j = 1, ..., n. \end{aligned} \tag{2.13}$$

These dynamics give a well-defined vector field on $\Re^{2n}$. (If $x_k = 0$, we take $r(\beta|x_k|)\frac{x_k^R w_{jk}^R + x_k^I w_{jk}^I}{|x_k|} = r(\beta|x_k|)\frac{x_k^I w_{jk}^R - x_k^R w_{jk}^I}{|x_k|} = 0$.) The equilibrium points of the dynamics are points where $\dot{x}_1^R = \dot{x}_1^I = ... = \dot{x}_n^R = \dot{x}_n^I = 0$. For purposes of the proof of convergence, the state space is taken to be $\Re^{2n}$.

At any point in the state space $\Re^{2n}$, except where x_j^R and x_j^I are both zero for some j, we can define a valid (nonsingular) change of coordinates by $x_j^R + ix_j^I = a_j e^{i\alpha_j}, \quad j = 1, ..., n$. In the new coordinates, the dynamics become:

$$\dot{a}_j = -a_j + \sum_k r(\beta a_k) b_{jk} \cos(\alpha_k - \alpha_j - \theta_{jk}) \tag{2.14}$$

$$\dot{\alpha}_j = \frac{1}{a_j} \sum_k r(\beta a_k) b_{jk} \sin(\alpha_k - \alpha_j - \theta_{jk}) \tag{2.15}$$

where $w_{jk}^R + i w_{jk}^I = b_{jk} e^{i\theta_{jk}}$.

At all points in $\Re^{2n}$ where the change of coordinates is valid, we define the Lyapunov function:

$$\begin{aligned} V &= -\sum_{j<k} r(\beta a_j) r(\beta a_k) b_{jk} \cos(\alpha_k - \alpha_j - \theta_{jk}) \\ &\quad -T \sum_j [-\beta a_j r(\beta a_j) + \log(2\pi I_0(\beta a_j))] \end{aligned} \tag{2.16}$$

where log denotes the natural log. (Keep in mind that $w_{jk} = w_{kj}^*$, $w_{jj} = 0$, and $\beta = \frac{1}{T}$.) It turns out that V can be continuously defined even where the change of coordinates is singular, because if either a_j or a_k is taken to be zero in the above formula for V, the term $r(\beta a_j) r(\beta a_k) b_{jk} \cos(\alpha_k - \alpha_j - \theta_{jk})$ will be zero regardless of the value of α_j or α_k.

Calculating $\dot{V}(\nu) = \frac{\partial V}{\partial \nu} \dot{\nu}$, we obtain

$$\begin{aligned} \dot{V}(\nu) &= -\sum_j \left\{ \beta r'(\beta a_j) \left[-a_j + \sum_k r(\beta a_k) b_{jk} \cos(\alpha_k - \alpha_j - \theta_{jk}) \right]^2 \right. \\ &\quad \left. + \frac{r(\beta a_j)}{a_j} \left[\sum_k r(\beta a_k) b_{jk} \sin(\alpha_k - \alpha_j - \theta_{jk}) \right]^2 \right\}. \end{aligned} \tag{2.17}$$

Note that $r(\beta a_j) > 0 \ \ \forall a_j > 0$ and $\frac{r(\beta a_j)}{a_j} > 0 \ \ \forall a_j > 0$. Also, $r'(\beta a_j) \to \frac{1}{2}$ as $a_j \to 0$, and $\frac{r(\beta a_j)}{a_j} \to \beta r'(\beta a_j)$ as $a_j \to 0$.

Thus, $\dot{V}(\nu) \leq 0 \ \ \forall a_j > 0 \ \ \forall \alpha_j$, and $\dot{V}(\nu) = 0$ if and only if

$$\begin{cases} -a_j + \sum_k r(\beta a_k) b_{jk} \cos(\alpha_k - \alpha_j - \theta_{jk}) = 0 \\ \sum_k r(\beta a_k) b_{jk} \sin(\alpha_k - \alpha_j - \theta_{jk}) = 0 \end{cases} \qquad \forall j = 1, ..., n. \tag{2.18}$$

But this will hold at a point $x_j = a_j e^{i\alpha_j} \ \ \forall j$ where the change of coordinates is valid if and only if it is an equilibrium point of the dynamics.

We now show that the Lyapunov function $V(\nu)$ has bounded sublevel sets, because this will enable us to exhibit compact sets which are positively invariant under the dynamics, as required to apply LaSalle's invariance principle. Observe that the first term of V, $-\sum_{j<k} r(\beta a_j) r(\beta a_k) b_{jk} \cos(\alpha_k - \alpha_j - \theta_{jk})$, is bounded as $a_j \to \infty$ for any (or all) a_j. A straightforward but lengthy calculation (relegated to Appendix A) shows that $\beta a_j r(\beta a_j) - \log(2\pi I_0(\beta a_j)) \to \infty \ \ as \ \ a_j \to \infty$. Then because $\beta a_j r(\beta a_j) - \log(2\pi I_0(\beta a_j)) > 0 \ \ \forall a_j > 0$, and because the terms $\beta a_j r(a_j) - \log(2\pi I_0(\beta a_j))$ appear summed in V, it follows immediately that V is radially unbounded in the a_j (where by definition $V : \Re^n \to \Re$ is radially unbounded in its argument $a \in \Re^n$ if $V(a) \to \infty$ as $||a|| \to \infty$ [17]). Moreover, since V is continuous even at points where the change of coordinates is singular, we can conclude that V has bounded sublevel sets when viewed as a function of $(x_1^R, x_1^I, ..., x_n^R, x_n^I)$.

So far we have shown that there is a Lyapunov function V continuous on all of $\Re^{2n}$, which has bounded sublevel sets, and which has $\dot{V}(\nu) < 0$

provided ν is not an equilibrium point and provided ν is not a point where our change of coordinates is singular. What we will now show is that there is no loss of generality in assuming that a trajectory will pass through points where the change of coordinates is singular only at isolated points in time.

Specifically, we will show is that if a trajectory has $x_j^R(t^*) = x_j^I(t^*) = 0$ for some $j \in \{1, ..., n\}$ and for some t^* then either t^* is an isolated point in time for which $x_j^R = x_j^I = 0$ or else $x_j^R = x_j^I = 0 \ \forall t > 0$. If $x_j^R = x_j^I = 0 \ \forall t > 0$, then simply eliminate the x_j^R and x_j^I coordinates and consider the reduced system of dimension $\Re^{2n-2}$. Repeating this test will reduce the system to a new system with coordinates $(\hat{x}_j^R, \hat{x}_j^I, ..., \hat{x}_{\hat{n}}^R, \hat{x}_{\hat{n}}^I)$ with $\hat{n} \leq n$. The trajectories of the reduced system will have $\hat{x}_j^R(t^*) = \hat{x}_j^I(t^*) = 0$ only at isolated time instants t^*, and hence we will have $\dot{V} = 0$ only at equilibrium points and at isolated times t^*. (It is easy to verify that the reduced system has exactly the same form in terms of dynamics, equilibria, and Lyapunov function as the original system.)

Analyticity properties are the key to showing that the system can be reduced so that $x_j^R = x_j^I = 0$ for some j only at isolated points in time. That the right-hand-side of the differential equation for $(x_1^R(t), x_1^I(t), ..., x_n^R(t), x_n^I(t))$, equation (2.13), is analytic in the variables $x_1^R, x_1^I, ..., x_n^R, x_n^I$ follows straightforwardly from basic properties of analyticity (using methods in, e.g., [18]).

This in turn implies that the trajectory $(x_1^R(t), x_1^I(t), ..., x_n^R(t), x_n^I(t))$ is an analytic function of t (the necessary results on analyticity and differential equations can be found in [19]). Therefore, for any $j \in \{1, ..., n\}$, (x_j^R, x_j^I) will be an analytic function of t. Hence if $x_j^R(t^*) = x_j^I(t^*) = 0$ then either t^* is an isolated point in time for which $x_j^R = x_j^I = 0$ or else $x_j^R = x_j^I = 0 \ \forall t > 0$.

We have a well-defined, C^∞ vector field (2.13) on $\Re^{2n}$ and a Lyapunov function V which has bounded sublevel sets. For simplicity, think of the vector field abstractly, as given by $\dot{\nu} = f(\nu)$, with Lyapunov function $V(\nu)$. Also, $\dot{V}(\nu) \leq 0$ along trajectories.

Fix $c > 0$ and let

$$\Omega_c = \{\nu \in \Re^{2n} | V(\nu) \leq c\}. \tag{2.19}$$

In order to apply LaSalle's Invariance Principle, we need to show that Ω_c is a compact positively invariant set. We have already shown that the Lyapunov function has bounded sublevel sets, and hence Ω_c is bounded. In fact, it can be easily shown that Ω_c is closed as well, and hence is compact. Positive invariance of Ω_c follows from the fact that $\dot{V}(\nu) \leq 0$ along trajectories.

Theorem 2.1 (LaSalle's Invariance Principle): Let Ω be a compact set and suppose the solution $\nu(t)$ starting in Ω stays in Ω for all $t > 0$. Let $V : \Omega \to \Re$ be a continuous function such that $V(\nu(t))$ is a monotone nonin-

creasing function of t. Let E be the set of all points in Ω where $\dot{V}(\nu)$ exists and equals zero. Let M be the largest positively invariant set in E. Then $\nu(t)$ approaches M as $t \to \infty$.

Proof: (See [17].)

Theorem 2.2: The dynamics (2.13) converge to an equilibrium point.

Proof: For any initial condition ν_0, reduce the system if necessary so that we may assume that the coordinate transformation is only singular at isolated points in time. Also, choose $c > 0$ to be greater than or equal to $V(\nu_0)$. Then the set Ω in Theorem 2.1 is taken to be Ω_c, as defined earlier, which is a well-defined compact subset of $\Re^{2n}$ when viewed in $(x_1^R, x_1^I, ..., x_n^R, x_n^I)$ coordinates. Because the vector field is well defined, the Lyapunov function is monotone nonincreasing, and the set Ω_c is compact, it follows that the trajectory $\nu(t)$ exists and stays in Ω_c $\forall t > 0$. Thus, the hypotheses of Theorem 2.1 are satisfied, enabling us to conclude that the trajectory will converge to the largest positively invariant subset of the set of points in Ω_c such that $\dot{V}(\nu) = 0$. But the positively invariant subset of the set of points with $\dot{V}(\nu) = 0$ are just the equilibrium points of the system which lie inside Ω_c.

LaSalle's Principle thus enables us to conclude convergence of any tra-

jectory to the set of equilibrium points of the dynamics, but not to a specific equilibrium point. We will now show, by appropriate choice of inner product, that the system follows gradient dynamics except at isolated points in time. At points $\nu = (a_1, ..., a_n, \alpha_1, ..., \alpha_n)$, define the inner product:

$$\langle \nu_1, \nu_2 \rangle = \nu_1^T \mathrm{diag}(\beta r'(\beta a_1), ..., \beta r'(\beta a_n), a_1 r(\beta a_1), ..., a_n r(\beta a_n)) \nu_2. \quad (2.20)$$

Then $\langle \dot{\nu}, \zeta \rangle = -d_\nu V \cdot \zeta = \langle -\nabla V, \zeta \rangle$ so that $\dot{\nu} = -\nabla V$; i.e., the system satisfies gradient dynamics (except at isolated time instants). Thus, we may conclude that in fact the dynamics (2.13) converge to an equilibium point. Q.E.D.

Observe that we can define two different but related sets of dynamics. As will be justified below, the following dynamics are more readily implementable:

$$\dot{y}_j = -y_j + r(\beta |x_j|) \frac{x_j}{|x_j|} \quad (2.21)$$

$$x_j = \sum_{k=1}^{n} y_k w_{jk}^* \quad (2.22)$$

$$\text{at equilibrium, } \; y_j = r(\beta |x_j|) \frac{x_j}{|x_j|}. \quad (2.23)$$

On the other hand, the following dynamics are more theoretically tractable:

$$\dot{x_j} = -x_j + \sum_k y_k w_{jk}^* \quad (2.24)$$

$$y_j = r(\beta |x_j|) \frac{x_j}{|x_j|} \quad (2.25)$$

$$\text{at equilibrium, } x_j = \sum_{k=1}^{n} y_k w_{jk}^* \tag{2.26}$$

(this is the network of Zemel et. al.)

The convergence result just proved applies directly to the second set of dynamics. For the first set of dynamics, the x_j variables satisfy the same differential equation as the x_j variables in the second set of dynamics. However, an additional argument (given in Appendix B) is needed to prove that the y_j variables in the first set of dynamics converge to $r(\beta|x_j|)\frac{x_j}{|x_j|}$ at equilibrium even if the weight matrix is singular.

The first set of dynamics is more readily implementable in an analog system (for instance an analog VLSI circuit, but the same argument would apply for a biological implementation) because there is a saturating summation followed by an integration (or low-pass filtering), with the state variables y_j constrained to lie inside the unit circle (for initial conditions inside the unit circle). By contrast, the second set of dynamics requires the storage and update of state variables taking values over a possibly much larger subset of the complex plane, with different x_j having possibly greatly different dynamic range requirements. On the other hand, the second set of dynamics is much easier to analyze, and the first step in the convergence proof was basically to convert the first set of dynamics into the second set of dynamics. A recurring theme in this work is that there are two ways of writing down

dynamics for what is essentially the same network (related by a change of variables). One version of the dynamics is more readily implemented, and the other is simpler to analyze mathematically.

2.4 Generalization to a Class of Networks

Although the choice of $r(\cdot)$ as the ratio of bessel functions enabled a physical motivation to be given for the dynamics, careful examination of the proof of convergence reveals that as long as $r(\cdot)$ satisfies certain properties, the convergence result will still hold. First, the function $r(\cdot)$ must be strictly monotone increasing with $r(0) = 0$. Second, $r(\cdot)$ must be analytic. Third, we need to ensure that a suitable radially unbounded Lyapunov function exists:

$$V = -\sum_{j<k} r(a_j)r(a_k)b_{jk}\cos(\alpha_k - \alpha_j - \theta_{jk}) + \sum_j g(a_j) \tag{2.27}$$

where $\lim_{a\to\infty} \frac{g(a)}{r^2(a)} = \infty$, $\frac{\partial g}{\partial a} = ar'(a)$, and $g(0) = \text{constant}$.

The ability to generalize the convergence proof to a class of networks in this manner is important when analog implementations of these networks are considered. Although saturating nonlinearities can be achieved in analog hardware, a saturating nonlinearity for the magnitude of a complex number which leaves the phase unaltered is more complicated.

Chapter 3

Adaptive Control Law for a Variable-Weight Network

3.1 Introduction

The usefulness of the network presented in the last section would be greatly enhanced if the weights were not constrained to be constant. An adaptive control law for variable weights is now described which retains the convergence properties of the fixed-weight network - namely, the variable weight network with adaptively controlled weights is shown to converge to an equilibrium point, again using LaSalle's Invariance Principle. The utility of the rigorous proof of convergence for the fixed-weight network is that it can be extended to more general and useful networks, of which the adaptive control law described in this chapter is an example.

3.2 Motivation for the Adaptive Control Law

We motivate the adaptive control law mathematically by viewing the dynamics and Lyapunov function abstractly and performing some calcula-

tions. Abstractly, for the fixed-weight case we have

$$\dot{\nu} = f(\nu, w) \text{ (dynamics)} \tag{3.1}$$

$$V(\nu, w) \qquad \text{(Lyapunov function)} \tag{3.2}$$

$$\begin{aligned} \dot{V}(\nu, w) &= \frac{\partial V}{\partial \nu} f(\nu, w) \\ &= - < f(\nu, w), f(\nu, w) > . \end{aligned} \tag{3.3}$$

Now letting some of the weights be feedback functions, we have

$$\dot{\nu} = f(\nu, w(\nu)) + g(\nu) \text{ (dynamics)} \tag{3.4}$$

$$V(\nu, w(\nu)) \qquad \text{(Lyapunov function)} \tag{3.5}$$

$$\begin{aligned} \dot{V}(\nu, w(\nu)) &= \left(\frac{\partial V}{\partial \nu} + \frac{\partial V}{\partial w} \frac{\partial w}{\partial \nu} \right) (f(\nu, w(\nu)) + g(\nu)) \\ &= - < f(\nu, w(\nu)) + g(\nu), f(\nu, w(\nu)) + g(\nu) > \end{aligned} \tag{3.6}$$

provided

$$\left(\frac{\partial V}{\partial w} \frac{\partial w}{\partial \nu} \right) \eta = - < g(\nu), \eta > . \tag{3.7}$$

In the transformed coordinates, this condition becomes

$$\begin{aligned} \frac{\partial V}{\partial w} \frac{\partial w}{\partial \nu} &= -[\beta r'(\beta a_1) g_1(\nu) \cdots \beta r'(\beta a_n) g_n(\nu) \\ &\qquad a_1 r(\beta a_1) g_{n+1}(\nu) \cdots a_n r(\beta a_n) g_{2n}(\nu)]. \end{aligned} \tag{3.8}$$

Suppose unit j interconnects units l_1 and l_2, and all other weights are constants. Furthermore, suppose b_{l_1,l_2} depends only on a_j and θ_{l_1,l_2} depends

only on α_j. Then

$$\begin{aligned}\frac{\partial V}{\partial w}\frac{\partial w}{\partial \nu} &= \Big[0\cdots 0 \;\; -r(\beta a_{l_1})r(\beta a_{l_2})\frac{\partial b_{l_1,l_2}}{\partial a_j}\cos(\alpha_{l_2}-\alpha_{l_1}-\theta_{l_1,l_2}) \;\; 0\cdots 0 \\ &\qquad 0\cdots 0 \;\; -r(\beta a_{l_1})r(\beta a_{l_2})b_{l_1,l_2}\sin(\alpha_{l_2}-\alpha_{l_1}-\theta_{l_1,l_2})\frac{\partial \theta_{l_1,l_2}}{\partial \alpha_j} \;\; 0\cdots 0\Big]\end{aligned} \tag{3.9}$$

where the first nonzero term is in position j of the row vector, and the second nonzero term is in position $j+n$.

If we choose $b_{l_1,l_2}(a_j) = d_{l_1,l_2}r(\beta a_j)$ and $\theta_{l_1,l_2}(\alpha_j) = \alpha_j$, then $\frac{\partial b_{l_1,l_2}}{\partial a_j} = d_{l_1,l_2}\beta r'(\beta a_j)$ and $\frac{\partial \theta_{l_1,l_2}}{\partial \alpha_j} = 1$, so

$$\begin{aligned}\frac{\partial V}{\partial w}\frac{\partial w}{\partial \nu} &= -\Big[0\cdots 0 \;\; \beta r'(\beta a_j)(d_{l_1,l_2}r(\beta a_{l_1})r(\beta a_{l_2})\cos(\alpha_{l_2}-\alpha_{l_1}-\alpha_j)) \;\; 0\cdots 0 \\ &\qquad 0\cdots 0 \; a_j r(\beta a_j)\left(\frac{1}{a_j}d_{l_1,l_2}r(\beta a_{l_1})r(\beta a_{l_2})\sin(\alpha_{l_2}-\alpha_{l_1}-\alpha_j)\right) \;\; 0\cdots 0\Big]\end{aligned} \tag{3.10}$$

where again the nonzero terms are in positions j and $j+n$ of the row vector.

Thus, letting

$$\begin{aligned} g_j(\nu) &= d_{l_1,l_2}r(\beta a_{l_1})r(\beta a_{l_2})\cos(\alpha_{l_2}-\alpha_{l_1}-\alpha_j) \\ g_{j+n}(\nu) &= \frac{1}{a_j}d_{l_1,l_2}r(\beta a_{l_1})r(\beta a_{l_2})\sin(\alpha_{l_2}-\alpha_{l_1}-\alpha_j) \\ g_k(\nu) &= 0, \quad k \neq j, j+n \end{aligned} \tag{3.11}$$

we then have

$$(\frac{\partial V}{\partial w}\frac{\partial w}{\partial \nu})\eta = - < g(\nu), \eta > . \tag{3.12}$$

Generalizing to an arbitrary number of variable weights, we have

$$I^j_{l_1,l_2} = \begin{cases} 1 & : \quad b_{l_1,l_2}e^{i\theta_{l_1,l_2}} = d_{l_1,l_2}r(\beta a_j)e^{i\alpha_j} \\ 0 & : \quad \text{otherwise} \end{cases} \tag{3.13}$$

$$\dot{a}_j = -a_j + \sum_k r(\beta a_k)b_{jk}\cos(\alpha_k - \alpha_j - \theta_{jk}) + \sum_{l_1,l_2} I^j_{l_1,l_2} d_{l_1,l_2} r(\beta a_{l_1}) r(\beta a_{l_2}) \cos(\alpha_{l_2} - \alpha_{l_1} - \alpha_j) \tag{3.14}$$

$$\dot{\alpha}_j = \frac{1}{a_j}\Big(\sum_k r(\beta a_k)b_{jk}\sin(\alpha_k - \alpha_j - \theta_{jk}) + \sum_{l_1,l_2} I^j_{l_1,l_2} d_{l_1,l_2} r(\beta a_{l_1}) r(\beta a_{l_2}) \sin(\alpha_{l_2} - \alpha_{l_1} - \alpha_j)\Big) \tag{3.15}$$

which (provided the change of coordinates is nonsingular) is equivalent to

$$I^j_{l_1,l_2} = \begin{cases} 1 & : \quad w_{l_1,l_2} = d_{l_1,l_2}r(\beta|x_j|)\frac{x_j}{|x_j|} \\ 0 & : \quad \text{otherwise} \end{cases} \tag{3.16}$$

$$\dot{x}_j = -x_j + \sum_k r(\beta|x_k|)\frac{x_k}{|x_k|}w^*_{jk} + \sum_{l_1,l_2} I^j_{l_1,l_2} d_{l_1,l_2} r(\beta|x_{l_1}|) r(\beta|x_{l_2}|) \frac{x_{l_2}}{|x_{l_2}|}\frac{x^*_{l_1}}{|x_{l_1}|} \tag{3.17}$$

where the conditions $b_{l_1,l_2}e^{i\theta_{l_1,l_2}} = d_{l_1,l_2}r(\beta a_j)e^{i\alpha_j}$ and $w_{l_1,l_2} = d_{l_1,l_2}r(\beta|x_j|)\frac{x_j}{|x_j|}$ simply indicate that unit j is serving as the interconnecting weight between units l_1 and l_2, with the sense of the connection (recall that $w_{kj} = w^*_{jk}$) taken into account.

3.3 Proof of Convergence for the Adaptive Control Law

To prove convergence of the adaptive control law, we start with the dy-

namical equations just given, (3.13)-(3.17), along with the Lyapunov function

$$\begin{aligned} V &= -\sum_{j<k} r(\beta a_j) r(\beta a_k) b_{jk} \cos(\alpha_k - \alpha_j - \theta_{jk}) \\ &\quad -T \sum_j \left[-\beta a_j r(\beta a_j) + \log\left(2\pi I_0(\beta a_j)\right) \right] \end{aligned} \tag{3.18}$$

where now we may have $b_{jk} = d_{jk} r(\beta a_l)$ and $\theta_{jk} = \pm\alpha_l$ for various j, k, l.

The new calculation of $\dot{V}$ gives:

$$\begin{aligned} \dot{V}(\nu) &= -\sum_j \Bigg\{ \beta r'(\beta a_j) \Bigg[-a_j + \sum_k r(\beta a_k) b_{jk} \cos(\alpha_k - \alpha_j - \theta_{jk}) \\ &\quad + \sum_{l_1, l_2} I^j_{l_1, l_2} d_{l_1, l_2} r(\beta a_{l_1}) r(\beta a_{l_2}) \cos(\alpha_{l_2} - \alpha_{l_1} - \alpha_j) \Bigg]^2 \\ &\quad + \frac{r(\beta a_j)}{a_j} \Bigg[\sum_k r(\beta a_k) b_{jk} \sin(\alpha_k - \alpha_j - \theta_{jk}) \\ &\quad + \sum_{l_1, l_2} I^j_{l_1, l_2} d_{l_1, l_2} r(\beta a_{l_1}) r(\beta a_{l_2}) \sin(\alpha_{l_2} - \alpha_{l_1} - \alpha_j) \Bigg]^2 \Bigg\}. \end{aligned} \tag{3.19}$$

From this point on, the proof of convergence is basically the same as the convergence proof for the fixed-weight case. The reason for labeling the feedback law an adaptive control law is that the original dynamics were linear in certain parameters (the weights), which are now adapted according to a feedback law which guarantees convergence. In this way, coupled oscillator networks can be designed to adapt in the presence of, for example, fixed but (a priori) unknown weight values, as the 6-unit feedback network presented in the next chapter will illustrate.

The implementable version of these dynamics is derived in Appendix

C. The implementable dynamics are

$$\dot{y}_j = -y_j + r(\beta|x_j|)\frac{x_j}{|x_j|} \tag{3.20}$$

$$x_j = \sum_{k=1}^{n} y_k w^*_{jk} + W_j - \sum_{l_1,l_2} I^{l_1}_{j,l_2} d_{j,l_2} y_{l_2} r(\beta|x_{l_1}|)\frac{x^*_{l_1}}{|x_{l_1}|} \tag{3.21}$$

$$\dot{W}_j = -W_j + \sum_{l_1,l_2} \left(I^j_{l_1,l_2} d_{l_1,l_2} + I^{l_1}_{j,l_2} d_{j,l_2} \right) r(\beta|x_{l_1}|) r(\beta|x_{l_2}|)\frac{x_{l_2}}{|x_{l_2}|}\frac{x^*_{l_1}}{|x_{l_1}|} \tag{3.22}$$

$$\text{at equilibrium, } y_j = r(\beta|x_j|)\frac{x_j}{|x_j|}. \tag{3.23}$$

The theoretically tractable dynamics are

$$\dot{x_j} = -x_j + \sum_k y_k w^*_{jk} + \sum_{l_1,l_2} I^j_{l_1,l_2} d_{l_1,l_2} y_{l_2} y^*_{l_1} \tag{3.24}$$

$$y_j = r(\beta|x_j|)\frac{x_j}{|x_j|} \tag{3.25}$$

$$\text{at equilibrium, } x_j = \sum_{k=1}^{n} y_k w^*_{jk} + \sum_{l_1,l_2} I^j_{l_1,l_2} d_{l_1,l_2} y_{l_2} y^*_{l_1}. \tag{3.26}$$

As was the case for the fixed-weight network, the implementable and theoretically tractable dynamics share the same convergence properties. The implementable dynamics still consist of a saturating summation followed by a low-pass filtering, except that now some of the terms which are summed by the saturating summer are low-pass filtered first. Furthermore, the function $r(\cdot)$ may be generalized in the same way as for the fixed-weight network, which in turn gives a convergent adaptive control law for a class of networks of nonlinear coupled oscillators.

Chapter 4

Eliminating Undesired Stable Equilibria

4.1 Introduction

To illustrate how coupled oscillator networks may be employed, and to motivate why it is necessary to devise a method of eliminating undesired stable equilibria, we now present and begin the analysis of the simple 6-unit feedback circuit shown below:

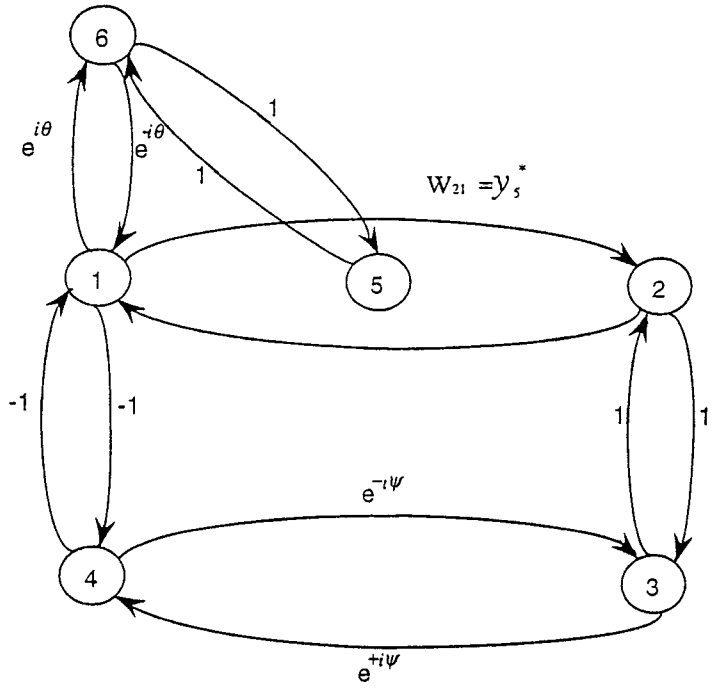


Figure 4.1: 6-unit feedback circuit.

θ represents an unkown phase shift between units 1 and 6, and as will be shown, the circuit acts to align the phase of unit 6 with $\pi + \psi$, a reference phase. Unit 5 interconnects units 1 and 2, and obeys the adaptive control law.

Since we know the network will converge to an equilibrium point, the first step is to find all of the equilibrium points and determine which are stable. The equilibrium equations in transformed coordinates are

$$\begin{aligned}
\dot{a}_1 &= -a_1 + r(\beta a_2) r(\beta a_5) \cos(\alpha_2 - \alpha_1 - \alpha_5) + r(\beta a_4) \cos(\alpha_4 - \alpha_1 - \pi) \\
&\quad + r(\beta a_6) \cos(\alpha_6 - \alpha_1 + \theta) = 0 \\
\dot{a}_2 &= -a_2 + r(\beta a_1) r(\beta a_5) \cos(\alpha_1 - \alpha_2 + \alpha_5) + r(\beta a_3) \cos(\alpha_3 - \alpha_2) = 0 \\
\dot{a}_3 &= -a_3 + r(\beta a_2) \cos(\alpha_2 - \alpha_3) + r(\beta a_4) \cos(\alpha_4 - \alpha_3 + \psi) = 0 \\
\dot{a}_4 &= -a_4 + r(\beta a_1) \cos(\alpha_1 - \alpha_4 - \pi) + r(\beta a_3) \cos(\alpha_3 - \alpha_4 - \psi) = 0 \\
\dot{a}_5 &= -a_5 + r(\beta a_6) \cos(\alpha_6 - \alpha_5) + r(\beta a_2) r(\beta a_1) \cos(\alpha_2 - \alpha_1 - \alpha_5) = 0 \\
\dot{a}_6 &= -a_6 + r(\beta a_1) \cos(\alpha_1 - \alpha_6 - \theta) + r(\beta a_5) \cos(\alpha_5 - \alpha_6) = 0 \\
\dot{\alpha}_1 &= \frac{1}{a_1} [r(\beta a_2) r(\beta a_5) \sin(\alpha_2 - \alpha_1 - \alpha_5) + r(\beta a_4) \sin(\alpha_4 - \alpha_1 - \pi) \\
&\quad + r(\beta a_6) \sin(\alpha_6 - \alpha_1 + \theta)] = 0 \\
\dot{\alpha}_2 &= \frac{1}{a_2} [r(\beta a_1) r(\beta a_5) \sin(\alpha_1 - \alpha_2 + \alpha_5) + r(\beta a_3) \sin(\alpha_3 - \alpha_2)] = 0 \\
\dot{\alpha}_3 &= \frac{1}{a_3} [r(\beta a_2) \sin(\alpha_2 - \alpha_3) + r(\beta a_4) \sin(\alpha_4 - \alpha_3 + \psi)] = 0 \\
\dot{\alpha}_4 &= \frac{1}{a_4} [r(\beta a_1) \sin(\alpha_1 - \alpha_4 - \pi) + r(\beta a_3) \sin(\alpha_3 - \alpha_4 - \psi)] = 0
\end{aligned}$$

$$
\begin{aligned}
\dot{\alpha}_5 &= \frac{1}{a_5}\left[r(\beta a_6)\sin(\alpha_6-\alpha_5)+r(\beta a_2)r(\beta a_1)\sin(\alpha_2-\alpha_1-\alpha_5)\right]=0 \\
\dot{\alpha}_6 &= \frac{1}{a_6}\left[r(\beta a_1)\sin(\alpha_1-\alpha_6-\theta)+r(\beta a_5)\sin(\alpha_5-\alpha_6)\right]=0.
\end{aligned}
\tag{4.1}
$$

The equilibrium phases we would like to show to be the only stable phases satisfy

$$
\begin{aligned}
\alpha_2-\alpha_1-\alpha_5 &= 0 \\
\alpha_4-\alpha_1-\pi &= 0 \\
\alpha_6-\alpha_1+\theta &= 0 \\
\alpha_3-\alpha_2 &= 0 \\
\alpha_4-\alpha_3+\psi &= 0 \\
\alpha_6-\alpha_5 &= 0.
\end{aligned}
\tag{4.2}
$$

The solution to this system of linear equations is

$$
\begin{aligned}
\alpha_1 &= \theta+\psi+\pi \\
\alpha_2 &= \theta+2\psi \\
\alpha_3 &= \theta+2\psi \\
\alpha_4 &= \theta+\psi \\
\alpha_5 &= \psi+\pi \\
\alpha_6 &= \psi+\pi.
\end{aligned}
\tag{4.3}
$$

However, although we would like to proceed with the equilibrium point identification and stability analysis in the transformed coordinates, there are two problems: first, there may be equilibria with one or more of the a_j equal to zero; and second, identifying all the equilibria (which means solving the system of nonlinear algebraic equations) is not a straightforward problem. Thus there may be undesirable stable equilibria in addition to the desired stable equilibrium point.

There are two distinct approaches one could take to eliminate undesired stable equilibria. The first approach is to use extra directional units configured to prevent the original network of units from having equilibrium points with one or more of the a_j being zero. We will show that by coupling one extra unit to each unit in the original network we can prevent any of the a_j from being zero at a stable equilibrium point. The second approach is to modify the network dynamics so that the a_j are constrained from being zero at equilibrium without the need for extra units. We will show that by simply assigning a value to the magnitude of a unit and eliminating the equation for that unit's magnitude from the dynamics, the same form of Lyapunov function still works for proving convergence. Roughly speaking, the first approach (adding units) is more implementable for an analog system, while the second approach (modifying the dynamical equations) is simpler to imple-

ment in a digital system. Again, the arguments used in the rigorous proof of convergence for the fixed-weight network will be instrumental in solving this technical problem, which must be overcome if such networks are to be useful in practice.

4.2 Digital Implementation Approach

The new dynamical equations, with the $a_j(t) = \tilde{a}_j$ constant functions of time, are

$$\begin{aligned} \dot{\alpha}_j &= \frac{1}{\tilde{a}_j}\Big(\sum_k r(\beta\tilde{a}_k) b_{jk} \sin(\alpha_k - \alpha_j - \theta_{jk}) \\ &\qquad + \sum_{l_1,l_2} I^j_{l_1,l_2} d_{l_1,l_2} r(\beta\tilde{a}_{l_1}) r(\beta\tilde{a}_{l_2}) \sin(\alpha_{l_2} - \alpha_{l_1} - \alpha_j)\Big). \end{aligned} \tag{4.4}$$

The Lyapunov function has the usual form:

$$\begin{aligned} V &= -\sum_{j<k} r(\beta\tilde{a}_j) r(\beta\tilde{a}_k) b_{jk} \cos(\alpha_k - \alpha_j - \theta_{jk}) \\ &\qquad -T\sum_j \left[-\beta\tilde{a}_j r(\beta\tilde{a}_j) + \log\left(2\pi I_0(\beta\tilde{a}_j)\right)\right]. \end{aligned} \tag{4.5}$$

Because the term $-T\sum_j \left[-\beta\tilde{a}_j r(\beta\tilde{a}_j) + \log\left(2\pi I_0(\beta\tilde{a}_j)\right)\right]$ is constant, we can just as well use the Lyapunov function

$$V = -\sum_{j<k} r(\beta\tilde{a}_j) r(\beta\tilde{a}_k) b_{jk} \cos(\alpha_k - \alpha_j - \theta_{jk}). \tag{4.6}$$

The convergence proof for this network is simpler than the general adaptive control law proof, because the system is evolving on the n-torus, T^n, rather than on $\mathbb{C}^n$. Ω_c is now simply T^n, which is a compact manifold. $\dot{V}(\nu)$ for $\nu = (\alpha_1, ..., \alpha_n)$ is computed to be

$$\begin{aligned}\dot{V}(\nu) &= -\sum_j \left\{ \frac{r(\beta a_j)}{a_j} \Big[\sum_k r(\beta a_k) b_{jk} \sin(\alpha_k - \alpha_j - \theta_{jk}) \right. \\ &\quad \left. + \sum_{l_1, l_2} I^j_{l_1, l_2} d_{l_1, l_2} r(\beta a_{l_1}) r(\beta a_{l_2}) \sin(\alpha_{l_2} - \alpha_{l_1} - \alpha_j) \Big]^2 \right\}. \quad (4.7)\end{aligned}$$

Thus, $\dot{V} \leq 0$ and $\dot{V} = 0$ only at equilibrium points of the dynamics. Applying LaSalle's Invariance Principle, we may conclude convergence to the set of equilibrium points. Furthermore, the dynamics are gradient dynamics, as we may define the inner product

$$\langle \nu_1, \nu_2 \rangle = \nu_1^T \mathrm{diag}(\tilde{a}_1 r(\beta \tilde{a}_1), ..., \tilde{a}_n r(\beta \tilde{a}_n)) \nu_2. \quad (4.8)$$

As before, $\langle \dot{\nu}, \zeta \rangle = -d_\nu V \cdot \zeta = \langle -\nabla V, \zeta \rangle$ so that $\dot{\nu} = -\nabla V$. Thus, we may conclude that the system actually converges to an equilibrium point.

Observe that the dynamics (4.4) are readily implementable in digital hardware, because the α_j and $\dot{\alpha}_j$ remain bounded (with a bound which is easily computed). Furthermore, $r(\cdot)$ only appears as its values at $\beta\tilde{a_1}, ..., \beta\tilde{a_n}$.

It may not always be desirable to prevent every unit in the network from being able to approach zero. The above proof can easily be extended to

accomodate having certain units' magnitudes fixed, but allowing other units' magnitudes to vary according to the adaptive control law.

4.3 Analog Implementation Approach

The dynamics (4.4), although easily implemented in digital hardware, are poorly suited to analog implementation due to the need to calculate the sine of various quantities. An alternative approach to modifying the dynamical equations (3.20) - (3.26) in order to eliminate the undesired stable equilibria is to instead modify the network topology. As will now be shown, by coupling an extra unit to each unit in the original network, stable equilibria with one or more of the a_j equal to 0 can be eliminated.

To illustrate the idea behind the approach of eliminating undesired stable equilibria by coupling an extra unit to each unit in the network, consider the following two-unit network:

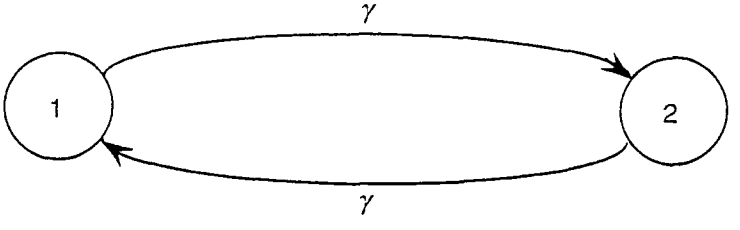


Figure 4.2: 2-unit network.

The complex dynamics for this system are

$$\dot{x}_1 = -x_1 + r(\beta|x_2|)\frac{x_2}{|x_2|}\gamma \tag{4.9}$$

$$\dot{x}_2 = -x_2 + r(\beta|x_1|)\frac{x_1}{|x_1|}\gamma. \tag{4.10}$$

We can show that this system is unstable at the origin by linearization. The Jacobian (found by expanding I_0 and I_1 about zero) is

$$\begin{bmatrix} -1 & \frac{\beta\gamma}{2} \\ \frac{\beta\gamma}{2} & -1 \end{bmatrix}. \tag{4.11}$$

Thus, the origin will be unstable provided $\beta\gamma > 2$.

Now we consider unit 1 to be connected to other units, and in addition unit 1 is connected to an auxiliary unit, unit 1', as shown:

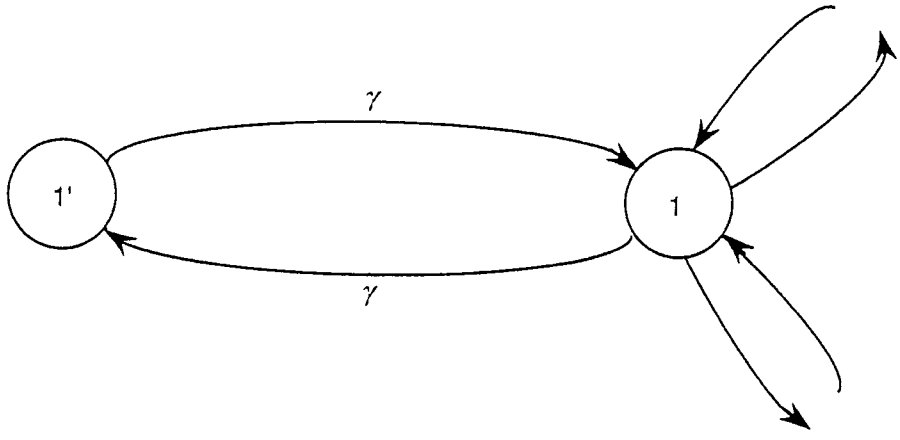


Figure 4.3: Unit connected to auxiliary unit and to rest of network.

The equilibrium equations for units 1 and 1' are

$$-x_1 + \sum_{k \neq 1'} r(\beta|x_k|)\frac{x_k}{|x_k|}w^*_{1k} + r(\beta|x_{1'}|)\frac{x_{1'}}{|x_{1'}|}\gamma = 0 \quad (4.12)$$

$$-x_{1'} + r(\beta|x_1|)\frac{x_1}{|x_1|}\gamma = 0. \quad (4.13)$$

If $\sum_{k \neq 1'} r(\beta|x_k|)\frac{x_k}{|x_k|}w^*_{1k} \neq 0$ then clearly the equilibrium equations cannot be satisfied if either $x_1 = 0$ or $x_{1'} = 0$. On the other hand, if $\sum_{k \neq 1'} r(\beta|x_k|)\frac{x_k}{|x_k|}w^*_{1k} = 0$ at equilibrium and we let $x_1 = x_{1'} = 0$, then we have an unstable linear system subject to a nonlinear perturbation which converges to zero. Therefore, we may conclude that the equilibrium point $x_1 = x_{1'} = 0$ is still unstable.

Since at no stable equilibrium point do we have $x_1 = 0$ or $x_{1'} = 0$, we can perform the rest of the analysis in transformed coordinates. The dynamics for $\alpha_{1'}$ is

$$\dot{\alpha}_{1'} = \frac{1}{a_{1'}}\gamma r(\beta a_1)\sin(\alpha_1 - \alpha_{1'}). \quad (4.14)$$

At equilibrium, this implies

$$\sin(\alpha_1 - \alpha_{1'}) = 0. \quad (4.15)$$

We would like to show that the only stable solution is $\alpha_1 - \alpha_{1'} = 0$. However, this will be a consequence of the following result, which can also be extended to show that for the system of figure 4.1 with an additional unit coupled to each unit shown, the only stable equilibrium point is the desired stable equilibrium point:

Theorem 4.1: (Result on stability of equilibrium points for fixed-weight networks) If the following conditions are satisfied:

(1) The network consists of interconnected units and is such that no stable equilibria exist with one or more $a_j = 0$.

(2) The network is connected; i.e., the network cannot be separated into two or more subnetworks whose convergence could be analyzed separately.

(3) All equilibrium points satisfy $\sin(\alpha_k - \alpha_j - \theta_{jk}) = 0 \quad \forall j, k$ such that $b_{jk} \neq 0$.

(4) The interconnecting weights are fixed.

Then the unique stable equilibrium point satisfies $\cos(\alpha_k - \alpha_j - \theta_{jk}) = 1 \quad \forall j, k$ such that $b_{jk} \neq 0$.

Proof: Suppose for an n-unit fixed-weight network for which no stable equilibria exist with one or more $a_j = 0, \quad \exists p, \ 1 \leq p < n$, such that at equilibrium

$$\cos(\alpha_k - \alpha_j - \theta_{jk}) = -1 \quad \text{for} \begin{cases} j = 1, 2, ..., p \\ k = p+1, ..., n \\ b_{jk} \neq 0 \end{cases} \tag{4.16}$$

$$\cos(\alpha_k - \alpha_l - \theta_{lk}) = 1 \quad \text{for} \begin{cases} l = p+1, ..., n \\ k = p+1, ..., n \\ b_{lk} \neq 0 \end{cases} \tag{4.17}$$

$$\cos(\alpha_m - \alpha_j - \theta_{jm}) \;=\; 1 \quad \text{for} \begin{cases} j = 1, 2, ..., p \\ m = 1, 2, ..., p \\ b_{jm} \neq 0. \end{cases} \tag{4.18}$$

(To see that this covers all possible cases for undesired stable equilibria, envision starting out with each unit satisfying the desired stable equilibrium configuration, and then changing the phase of certain units by π. This involves no loss of generality because for fixed-weight networks the equilibria are classified by phase differences rather than absolute phases. In other words, if a given set of phases represents an equilibrium point, adding a constant to each phase gives a different set of phases, but the new set also represents an equilibrium point with the same stability properties.)

Let

$$\hat{\alpha}_j = \alpha_j + S_j \Delta\alpha \tag{4.19}$$

where

$$S_j = \begin{cases} 1 & j = 1, ..., p \\ 0 & j = p+1, ..., n \end{cases} \tag{4.20}$$

and $0 < |\Delta\alpha| << 1$. Letting $\Delta V = V(\hat{\nu}) - V(\nu)$, we have

$$\begin{aligned} \Delta V \;&=\; \sum_{j<k} r(\beta a_j) r(\beta a_k) b_{jk} [-\cos((\alpha_k - \alpha_j - \theta_{jk}) + (S_k - S_j)\Delta\alpha) \\ &\quad + \cos(\alpha_k - \alpha_j - \theta_{jk})] \qquad (4.21) \\ &=\; \sum_{j<k} r(\beta a_j) r(\beta a_k) b_{jk} \cos(\alpha_k - \alpha_j - \theta_{jk}) \left[1 - \cos((S_k - S_j)\Delta\alpha)\right]. \qquad (4.22) \end{aligned}$$

Assumption (2) implies that the summation above must have at least one nonzero term. Furthermore, $[1-\cos((S_k - S_j)\Delta\alpha)] > 0$ for $j \leq p$ and $k > p$, which is precisely when $\cos(\alpha_k - \alpha_j - \theta_{jk}) = -1$. Otherwise, $[1 - \cos((S_k - S_j)\Delta\alpha)] = 0$. Thus, $\Delta V < 0$, which means there is a descent direction at any equilibrium point which we desire to be an unstable equilibrium point. Moreover, the dynamics in a neighborhood of one of these equilibrium points are gradient dynamics (assumption (1) guarantees the metric is well-defined in a neighborhood of the equilibria). Thus, all equilibria not satisfying $\cos(\alpha_k - \alpha_j - \theta_{jk}) = 1 \ \ \forall j, k$ such that $b_{jk} \neq 0$ must be unstable. Q.E.D.

Theorem 4.2: The network of figure 4.1, with an auxiliary unit connected to each unit shown to preclude stable equilibria with one or more $a_j = 0$, has the unique stable equilibrium point given by (4.3).

Proof: The proof of theorem 4.1 must be generalized slightly for the case at hand, in which one of the weights is nonconstant and is adaptively controlled. (4.16) - (4.18) still apply, exluding the case when w_{jk} is the adaptively controlled weight in (4.16), when w_{lk} is the adaptively controlled weight in (4.17), and when w_{jm} is the adaptively controlled weight in (4.18). The justification for using (the modified) (4.16) - (4.18) can no longer be based on the fact that for fixed-weight networks only the relative phases matter, be-

cause the absolute phases are important in the adaptively-controlled-weight case. Here we can use the fact that $\sin(\phi) = 0$ implies $\cos(\phi) = \pm 1$, and view (4.16) - (4.18) as simply an enumeration of which cosine terms are 1 and which are -1. (The relatively simple interconnection of units in this network makes it easy to see that this is possible.)

We define $\hat{\alpha}$ as in (4.19). However, S_j must now be chosen based on whether $\cos(\alpha_k - \alpha_j - \alpha_l) = 1$ or -1, where unit l is the adaptively controlled interconnecting weight for units j and k. The equation for ΔV is now

$$\Delta V = \sum_{j<k} r(\beta a_j) r(\beta a_k) b_{jk} \cos(\alpha_k - \alpha_j - \theta_{jk})[1 - \cos((S_k - S_j - S_{jk})\Delta\alpha)] \tag{4.23}$$

where $S_{jk} = S_l$ if unit l is the adaptively controlled weight and $S_{jk} = 0$ otherwise.

If $\cos(\alpha_k - \alpha_j - \alpha_l) = -1$, then S_j may be defined as in (4.20). However, if $\cos(\alpha_k - \alpha_j - \alpha_l) = 1$, then S_j must be chosen to make $S_k - S_j - S_l = 0$. S_j must also satisfy

$$S_j = \begin{cases} c_1 & j = 1, ..., p \\ c_2 & j = p+1, ..., n \end{cases} \tag{4.24}$$

where c_1 and c_2 are constants with $c_1 \neq c_2$. (Regardless of the values of c_1 and c_2, the fixed-weight terms of ΔV will be nonpositive.) Given whether j, k, and l are greater than or less than p, it is clearly possible to find c_1 and

c_2 so that $S_k - S_j - S_l = 0$ is satisfied.

(There is one final special case to take into account: the $p = 0$ case with $\cos(\alpha_k - \alpha_j - \alpha_l) = -1$. An appropriate function S_j giving $\Delta V < 0$ is readily obtainable in this case as well.)

Hence, as in the proof of Theorem 4.1, $\Delta V < 0$, which implies that all equilibria not satisfying $\cos(\alpha_k - \alpha_j - \theta_{jk}) = 1 \ \ \forall j, k$ such that $b_{jk} \neq 0$ must be unstable. Therefore, the unique stable equilibrium phases for the system of figure 4.1 are given by (4.3). Q.E.D.

Thus, as claimed, the circuit of figure 4.1 acts to align the phase of unit 6 with $\pi + \psi$, a reference phase, regardless of the unknown phase shift θ between units 1 and 6.

Chapter 5

Analog VLSI Implementation

5.1 Introduction

There are several reasons to consider attempting to implement these networks of coupled oscillators in analog VLSI hardware. Analog processing at the output of receivers for communications or radar, or at the output of smart sensors, can improve system performance if the analog processing can be made to work well enough. The principle limitation for analog continuous-time silicon CMOS circuits is generally threshold voltage mismatch between nominally identical transistors. Such threshold voltage mismatches, or offsets, will affect how these coupled oscillator networks perform. However, the advantage the network of coupled oscillators possesses is that by keeping the unit magnitudes fairly large (at least around stable equilibria), the signal levels within the coupled oscillator network can be maintained at a high enough level that the offsets only introduce a small relative error. Furthermore, expressing the dynamics in the implementable form (2.21)-(2.23) or (3.20)-(3.23) aids in analyzing offset effects, in contrast to the theoretically tractable form (2.24)-(2.26) or (3.24)-(3.26) which fails to indicate the

dynamic range requirements of the x_j.

Analog VLSI is the most realistic option for implementing these networks in an analog form, because the number of multipliers required is fairly large even for networks consisting of a modest number of units. Only silicon CMOS implementation will be considered in this work, because silicon CMOS is the most widely used and available technology for prototyping analog VLSI circuits. Furthermore, to make the discussion more concrete, the MOSIS 2 μm technology will be considered [20].

Processing nonuniformities cause slight variations in threshold voltage, even for nominally identical transistors located in close proximity. Based on the author's experience in testing analog CMOS circuits at the U.S. Naval Research Laboratory's Microelectronic Device Physics Section as part of the analog VLSI adaptive filter project [21][22][23][24], it was generally found that nominally identical transistors within a circuit (i.e., transistors not wire-bonded to package leads) and located in close proximity had threshold mismatches in the 2mV to 3mV range. However, transistors with gates bonded to package leads regularly exhibited threshold mismatches as high as 200mV. One possible conclusion from this observation is that during the wire-bonding process, charge gets trapped in the gate oxide, causing a threshold voltage shift. For design work, the threshold mismatch between nearby identical

transistors was taken to be 10mV, as recommended in [25], since nearly all threshold offsets would be expected to be below that value, and care was taken with respect to how transistor gates were brought to bonding pads.

The purpose of this chapter is to sketch out an approach for implementing networks of coupled oscillators in analog VLSI in order to illuminate important practical issues. Two methods of producing the complex saturating nonlinearity are presented, and the extra dynamics they introduce are discussed.

5.2 Complex Saturating Nonlinearities

5.2.1 Introduction

The first challenge in devising an analog implementation of either the fixed-weight dynamics (2.21)-(2.23) or the adaptively controlled dynamics (3.20)-(3.22) lies in how to effect the complex saturating nonlinearity $x_j \mapsto r(\beta|x_j|)\frac{x_j}{|x_j|}$. Two quite different approaches will be presented: a quadrature oscillator approach and an automatic gain control approach. In both cases, it will be shown that the complex saturating nonlinearity introduces a high-frequency pole into the dynamics, and that this pole will limit the

maximum corner frequency of the network low-pass filters (to a corner frequency well below the corner of the high-frequency pole). The situation is similar to voltage op-amp design, where the ultimate bandwidth limitation of a normal voltage op-amp is due to the high-frequency non-dominant pole rather than the low-frequency dominant pole [26]. The quadrature oscillator approach has the advantage of being conceptually simple, but the automatic gain control approach will be shown to require fewer components.

5.2.2 Quadrature Oscillator Approach

The quadrature oscillator approach for producing the complex saturating nonlinearity is shown in figure 5.1. A reference oscillator generates I (in-phase) and Q (quadrature) signals, which are 90 degrees out of phase. These high-frequency signals multiply the real and imaginary parts of each $y_k w_{jk}^*$ term, and summing the high-frequency signals produces a high-frequency signal whose phase contains the directional information present in the real and imaginary parts of $y_k w_{jk}^*$. For each unit j, the high-frequency signals $y_k w_{jk}^*$ are summed (to produce x_j) and passed through a real saturating nonlinearity. At the output of the real saturating nonlinearity, the high-frequency sum signal for each unit (which is now a distorted sine wave) is multiplied by I and Q again, and low-pass-filtered to give the real and imaginary parts

of $r(\beta x_j)\frac{x_j}{|x_j|}$.

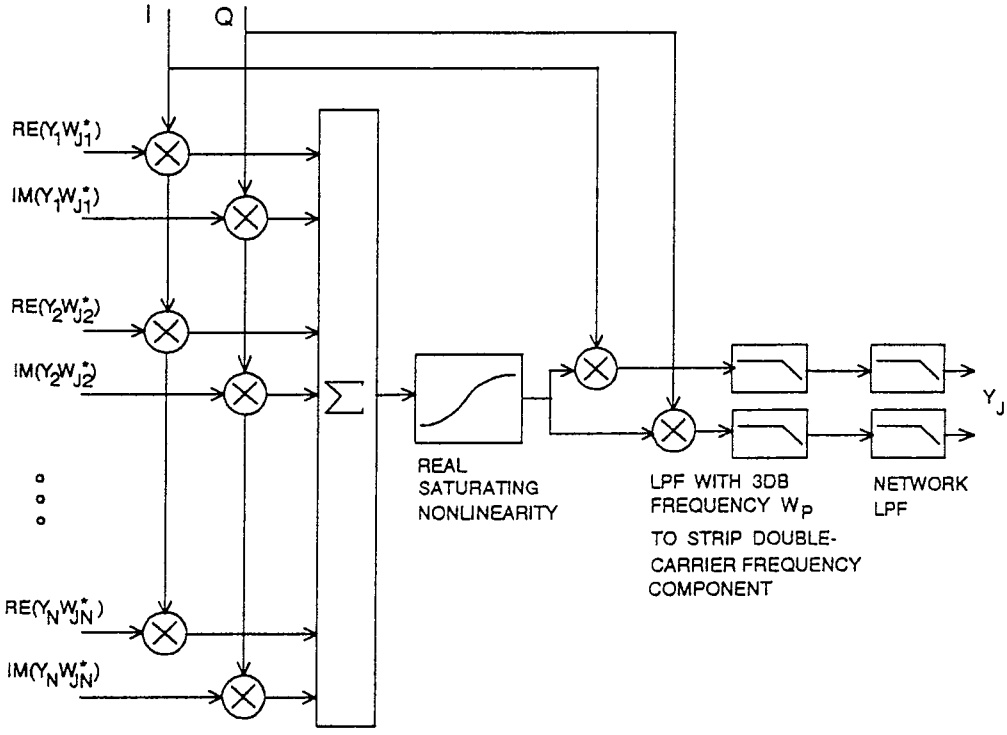


Figure 5.1: Quadrature oscillator approach.

This approach has the advantage of being conceptually simple, but there are two drawbacks. The first drawback is that the $y_k w_{jk}^*$ signals must be mixed up to the carrier frequency before they can be summed over k for each j. It is not possible to first sum the $y_k w_{jk}^*$ over k to produce the x_j and then mix the x_j up to the carrier frequency to be passed through the saturating nonlinearity, because the saturation operation must take place at

the point where the $y_k w_{jk}^*$ are summed. (Otherwise, the real and imaginary parts of $\sum_k y_k w_{jk}^*$ will saturate individually and the phase information in x_j will be lost.) As will be shown when actual circuits are considered, what are summed are currents, and the sum current can be converted to a voltage using a nonlinear resistance to produce the saturating nonlinearity.

The second and more subtle drawback is the need to properly handle the high-frequency signals as the $y_k w_{jk}^*$ are being summed and saturated to produce the $r(\beta x_j)\frac{x_j}{|x_j|}$. Any phase shift of the high-frequency signal as it is being passed through the saturating nonlinearity will result in a phase error once the $r(\beta x_j)\frac{x_j}{|x_j|}$ are demodulated. If too great, the high-frequency phase shift can cause the network of directional units to become unstable (or more specifically, to oscillate with a stable limit cycle solution). Therefore, there is a limit to how high the I-Q frequency can be for the network to still converge to a stable equilibrium point of the dynamics. Furthermore, since the high-frequency content of the $r(\beta x_j)\frac{x_j}{|x_j|}$ signal must be stripped so that it will not appear on y_j, the network low-pass-filter corner frequency must be significantly lower than the I-Q frequency. Either the network low-pass-filter corner frequency must be so low that there is sufficient attenuation of the high-frequency signal with only the 20dB/decade rolloff of a simple pole, or else a higher-frequency pole can be introduced to strip the double-carrier-

frequency component before the $r(\beta x_j)\frac{x_j}{|x_j|}$ signal is low-pass filtered. In either case, the network low-pass-filter corner frequency must be considerably lower than the bandwidth of the components used in the circuit. Even if no high-frequency pole is explicity introduced, there will still be high-frequency poles where the limitations of the multiplier and saturating nonlinearity bandwidth are reached.

5.2.3 Automatic Gain Control Approach

The automatic gain control approach for producing the complex saturating nonlinearity is shown in figure 5.2:

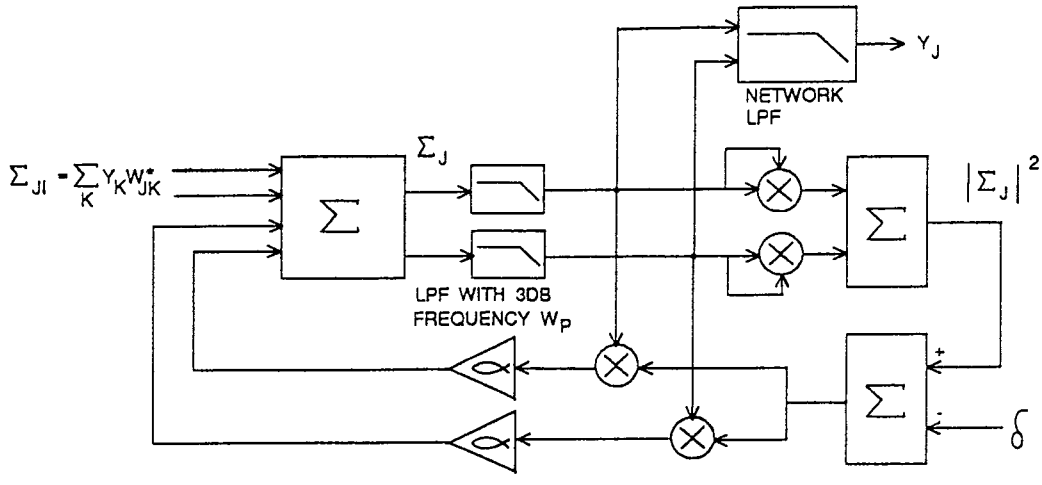


Figure 5.2: Automatic gain control approach.

It is a nonlinear feedback circuit acting like an automatic gain control for the signal entering the low-pass filter. The advantage of this approach over the quadrature oscillator approach is fewer components, since no oscillator and fewer multipliers (only four per unit) are required. However, some effort is needed to verify that the automatic gain control circuit serves as an appropriate complex saturating nonlinearity, and that as a subsystem it is stable in its own right.

The reason for choosing this particular form for the automatic gain control circuit is that it uses only multipliers, summers, and low-pass filters (the gain block labeled α can be incorporated into the other components). As will be shown, the low-pass filter in the feedback loop is required for the automatic gain control loop to remain stable. The location of the low-pass filter within the feedback loop has been chosen to help maintain the stability of the overall network of units.

The dc transfer characteristic of the automatic gain control circuit will now be shown to satisfy the requirements of the function $r(\cdot)$ enumerated in section 2.4. Referring to figure 5.2,

$$\Sigma_j = \Sigma_{jI} + \alpha(\delta - |\Sigma_j|^2)\Sigma_j \tag{5.1}$$

$$\Sigma_j(1 - \alpha\delta + \alpha|\Sigma_j|^2) = \Sigma_{jI} \tag{5.2}$$

$$|\Sigma_j|(1-\alpha\delta+\alpha|\Sigma_j|^2) = |\Sigma_{jI}|. \tag{5.3}$$

Therefore, we can write $r(a)$ as an implicit function of a:

$$r(a)(1-\alpha\delta+\alpha r^2(a)) = a \tag{5.4}$$

$$\alpha r^3(a)+(1-\alpha\delta)r(a)-a = 0 \tag{5.5}$$

$$r^3(a)+(\frac{1}{\alpha}-\delta)r(a)-\frac{a}{\alpha} = 0. \tag{5.6}$$

The above is a cubic polynomial, which can be solved as follows [27]: We assume that $(\frac{1}{\alpha}-\delta)>0$. Then since $-\frac{a}{\alpha}<0$, we can conclude that there is a unique real solution for $r(a)$. This real solution is given by

$$r(a)=\sqrt[3]{\frac{a}{2\alpha}+\sqrt{\frac{a^2}{4\alpha^2}+\frac{(\frac{1}{\alpha}-\delta)^3}{27}}}+\sqrt[3]{\frac{a}{2\alpha}-\sqrt{\frac{a^2}{4\alpha^2}+\frac{(\frac{1}{\alpha}-\delta)^3}{27}}}. \tag{5.7}$$

Using the binomial formula and the fact that the composition of analytic functions is analytic, it is easy to see that this function $r(\cdot)$ is analytic. Furthermore, $r(0)=0$ and $r(\cdot)$ is strictly monotone increasing. Also, we can take $\frac{\partial g}{\partial a}=ar'(a)$, with $g(0)=$ constant, and for large a approximate:

$$r(a)\approx\frac{a^{1/3}}{\alpha^{1/3}}. \tag{5.8}$$

We can then further approximate:

$$r'(a) \approx \frac{1}{3}\frac{a^{-2/3}}{\alpha^{1/3}} \tag{5.9}$$

$$\frac{\partial g}{\partial a} \approx \frac{1}{3}\frac{a^{1/3}}{\alpha^{1/3}} \tag{5.10}$$

$$g(a) \approx \frac{1}{4}\frac{a^{4/3}}{\alpha^{1/3}} \tag{5.11}$$

$$\frac{g(a)}{r^2(a)} \approx \frac{1}{4}\alpha^{1/3}a^{2/3}. \tag{5.12}$$

Thus, we may conclude that $\lim_{a\to\infty}\frac{g(a)}{r^2(a)} = \infty$, and hence that the function $r(\cdot)$ produced by the circuit of figure 5.2 indeed leads to a convergent network of nonlinear coupled oscillators.

Next we need to show that the automatic gain control subcircuit viewed as a nonlinear feedback circuit is stable. It turns out this can be shown by a LaSalle's Invariance Principle argument along the lines of the proof of Theorem 2.2. To simplify notation for the stability proof, the circuit of figure 5.2 is redrawn in figure 5.3, with u and z complex:

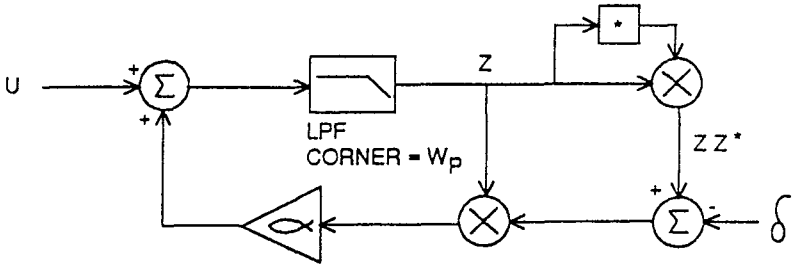


Figure 5.3: Automatic gain control circuit for stability proof.

Theorem 5.1: The circuit of figure 5.3 converges to the unique desired stable equilibrium point expressed in equation (5.7).

Proof: The complex dynamics are

$$\frac{1}{\omega_p}\dot{z} = -z + \alpha z(\delta - |z|^2) + u. \tag{5.13}$$

If we let $z = \rho e^{i\theta}$ and $u = \sigma e^{i\psi}$, we can rewrite the dynamics in polar coordinates as follows:

$$\begin{aligned}
\frac{1}{\omega_p}(\dot{\rho}e^{i\theta} + i\dot{\theta}\rho e^{i\theta}) &= -\rho e^{i\theta} + \alpha\rho e^{i\theta}(\delta - \rho^2) + \sigma e^{i(\psi-\theta)}e^{i\theta} && (5.14)\\
\frac{1}{\omega_p}(\dot{\rho} + i\dot{\theta}\rho) &= -\rho + \alpha\rho(\delta - \rho^2) \\
&\quad +\sigma(\cos(\psi-\theta) + i\sin(\psi-\theta)) && (5.15)\\
\frac{1}{\omega_p}\dot{\rho} &= -\rho + \alpha\rho(\delta - \rho^2) + \sigma\cos(\psi-\theta) && (5.16)\\
\frac{1}{\omega_p}\rho\dot{\theta} &= \sigma\sin(\psi-\theta). && (5.17)
\end{aligned}$$

Thus, provided $\rho \neq 0$, we can express the dynamics as

$$\begin{aligned}
\dot{\rho} &= \omega_p[-\rho + \alpha\rho(\delta - \rho^2) + \sigma\cos(\psi-\theta)] && (5.18)\\
\dot{\theta} &= \omega_p\frac{1}{\rho}[\sigma\sin(\psi-\theta)]. && (5.19)
\end{aligned}$$

The Lyapunov function is

$$V(\rho,\theta) = \frac{1}{2}\rho^2 - \alpha\left(\frac{\delta}{2}\rho^2 - \frac{1}{4}\rho^4\right) - \rho\sigma\cos(\psi-\theta). \tag{5.20}$$

The Lyapunov function is radially unbounded because of the $\frac{\alpha}{4}\rho^4$ term and therefore has compact sublevel sets. The derivative of the Lyapunov function along trajectories is computed as follows:

$$\dot{V}(\rho,\theta) = \frac{\partial V}{\partial\rho}\dot{\rho} + \frac{\partial V}{\partial\theta}\dot{\theta} \tag{5.21}$$

$$= -(-\rho + \alpha\rho(\delta - \rho^2) + \sigma\cos(\psi - \theta)\dot{\rho}$$

$$-(\rho\sigma\sin(\psi - \theta))\dot{\theta} \tag{5.22}$$

$$= -\omega_p[-\rho + \alpha\rho(\delta - \rho^2) + \sigma\cos(\psi - \theta)]^2$$

$$-\omega_p[\sigma\sin(\psi - \theta)]^2. \tag{5.23}$$

Thus, $\dot{V}(\rho, \theta) \leq 0$ and $\dot{V}(\rho, \theta) = 0$ only if (ρ, θ) is an equilibrium point of the dynamics, or else $\rho = 0$. But if $\rho = 0$, then examination of the original complex dynamics (5.13) reveals that the origin is only an invariant point of the dynamics if $u = 0$, in which case $\rho = 0$ is the unique desired stable equilibrium point. Otherwise, it is clear from the complex dynamics (5.13) and the assumption that $(\frac{1}{\alpha} - \delta) > 0$ that the unique equilibrium point of the system has $\theta = \psi$ and $\rho(\sigma)$ equivalent to $r(a)$ given in (5.7) for $r(\cdot) = \rho(\cdot)$ and $a = \sigma$.

So by Theorem 2.1 (LaSalle's Invariance Principle), and the fact that the dynamics can be written as gradiant dynamics, we may conclude that the system (5.13) converges to the unique desired equilibrium point. Q.E.D.

In practice, the circuit of figure 5.3 will have additional dynamics besides the low-pass filter shown, since the multipliers have finite bandwidth. If there is some frequency at which the feedback loop has an overall gain greater than unity but the phase shift through the loop exceeds 180 degrees, then

clearly the circuit will be prone to oscillate. Therefore, the pole frequency ω_p must be chosen low enough that the unmodeled dynamics do not cause the feedback loop to go unstable.

5.3 Summary

Two methods of performing the complex saturating nonlinearity function for an analog continuous-time implementation of the network of coupled oscillators have been presented. The first, the quadrature oscillator approach, is conceptually simple, but requires two multipliers per unit and two multipliers per interconnecting weight for each nonzero interconnecting weight in the network. The second approach, the automatic gain control circuit, requires effort to prove that it can serve its purpose, but it requires no oscillator and only four multipliers per unit, substantially fewer than the quadrature oscillator approach even for the relatively sparcely interconnected circuit of figure 4.1. Either approach introduces a non-dominant pole, either explicitly or implicitly, which in turn determines the maximum network low-pass-filter corner frequency for which the overall network of units will be stable. Because of the complicated nonlinear nature of the network of coupled oscillators with additional dynamics due to the limited bandwidths of the consituent components, a precise formula for determining the maximum

stable network low-pass-filter corner frequency would be at best very difficult to derive. Therefore, the choice of network low-pass filter corner frequency must be made conservatively, and it should at a minimum be verified that with a given dominant pole and non-dominant pole, each pair of connected units taken pairwise will not have an obvious stable limit cycle solution at any frequency.

Chapter 6

Analog VLSI Circuits

6.1 Introduction

The goal of this chapter is to combine the practical issues of analog VLSI implementation discussed in the last chapter with actual circuit building blocks in order to explore tradeoffs among network size, low-pass-filter corner frequency, and power dissipation. Using this information the decision can be made as to whether an analog VLSI coupled-oscillator network is worth considering for a given application.

As demonstrated in the last chapter, analog multipliers are required even for implementing the fixed-weight dynamics (2.21)-(2.23), because even if the fixed weights are implementated as fixed-gain amplifiers, multipliers are still needed for producing the complex saturating nonlinearities. Therefore, a multiplier circuit is described which could be used with either complex saturating nonlinearity approach described in the last chapter, as well as to multiply the y_k by either fixed or variable weights w^*_{jk}. Next, another required circuit building block, the low-pass filter, is considered, and finally

the network size, corner frequency, and power dissipation issues are discussed.

6.2 Analog Multiplier Design

Among the options for an analog integrated multiplier circuit, I have chosen to describe one which is relatively simple to design, simple to understand, and has a wide bandwidth: the wide-range Gilbert multiplier of figure 6.1:

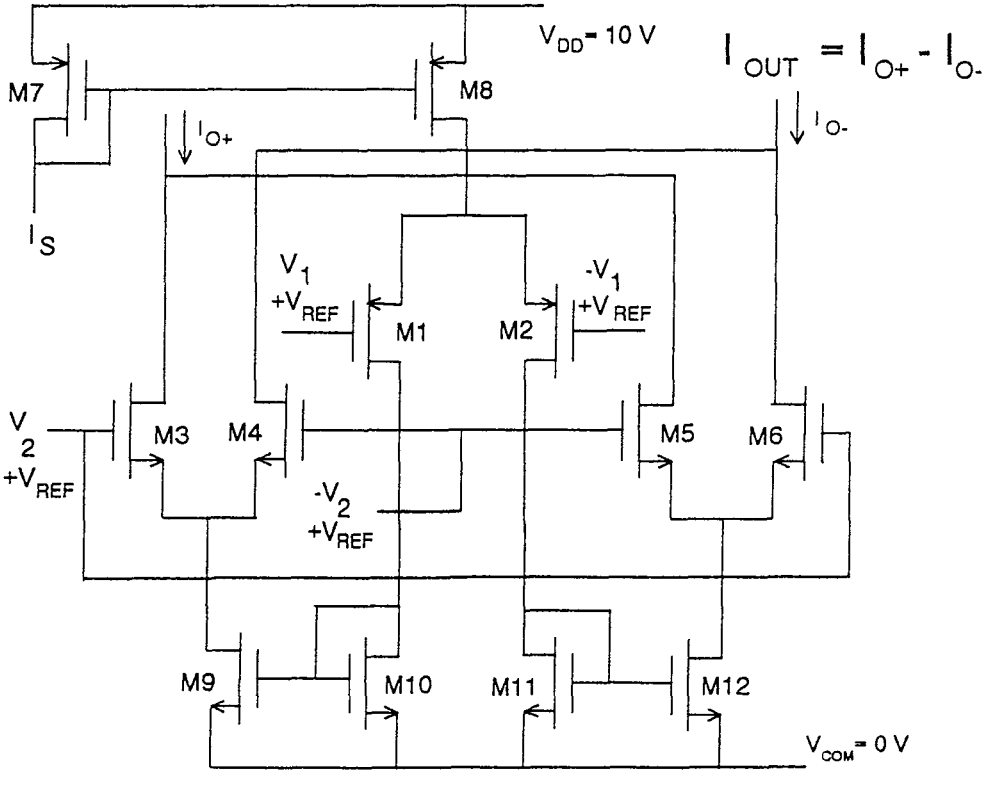


Figure 6.1: Wide-range Gilbert multiplier circuit.

This multiplier circuit has performed well in the adaptive filter project of the Microelectronic Device Physics Section at the Naval Research Laboratory [21],[23]. The wide-range Gilbert multiplier circuit has the feature that both differential inputs can be biased at the same dc voltage mid-range between the 0V and 10V power supplies, with a couple of volts of common-mode input range; therefore, the multiplier is easy to incorporate into a larger circuit. Furthermore, by adjusting the current and W/L values, a variety of input ranges, gains, and bandwidths can be obtained. The multiplier output is a differential current, which can be converted to a differential or single-ended voltage by using appropriate load devices (shown in figure 6.2), and the outputs of several multipliers can easily be summed simply by tying together the output current lines (as in figure 6.3).

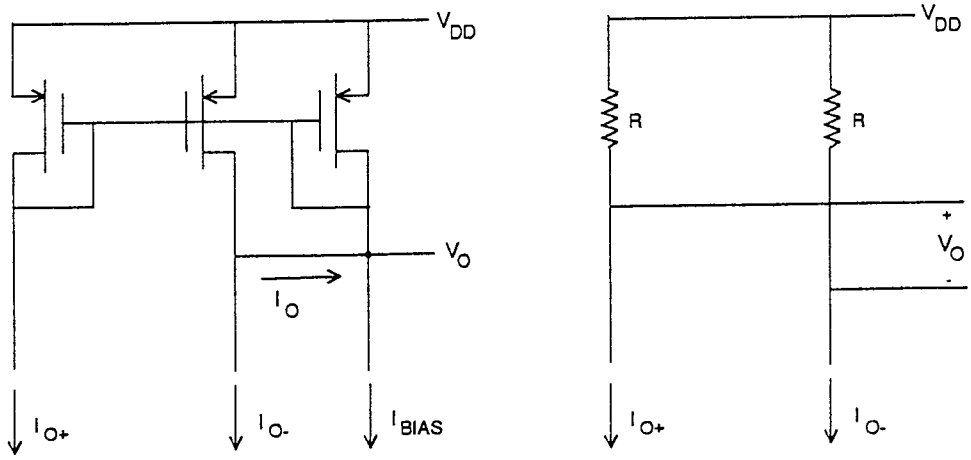


Figure 6.2: Multiplier load devices (left: single-ended, right: differential).

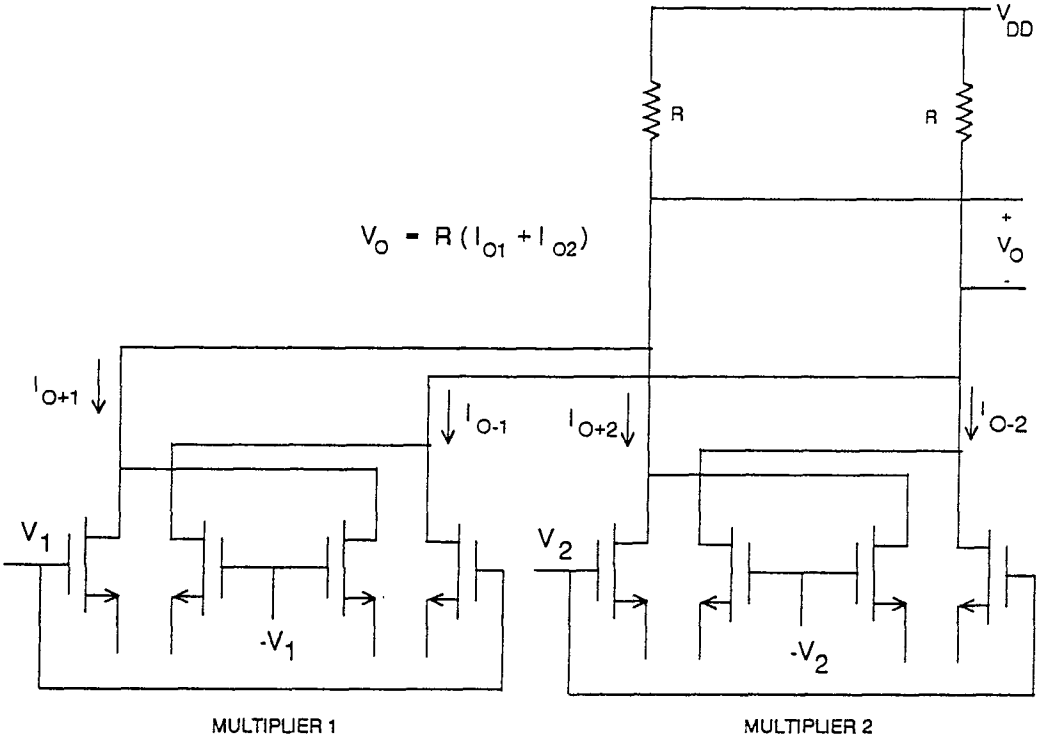


Figure 6.3: Summing the outputs of several multipliers.

To demonstrate how the circuit of figure 6.1 performs a multiplication operation, refer to figure 6.4 for the following first-order analysis of the standard Gilbert multiplier. All transistors are assumed to be in saturation, where the following (approximate) equation holds:

$$I_D = K(V_{GS} - V_T)^2 \tag{6.1}$$

where $K = \frac{1}{2}\mu C_{ox} W/L$, with μ the carrier mobility, C_{ox} the oxide capacitance per unit area, W the channel width, and L the channel length. I_D is the drain current, V_{GS} is the gate-source voltage, and V_T is the threshold voltage for

turning on the transistor. Transistors M1 and M2 are assumed to be identical with $K = K_1$, and transistors M3 through M6 are assumed to be identical with $K = K_2$.

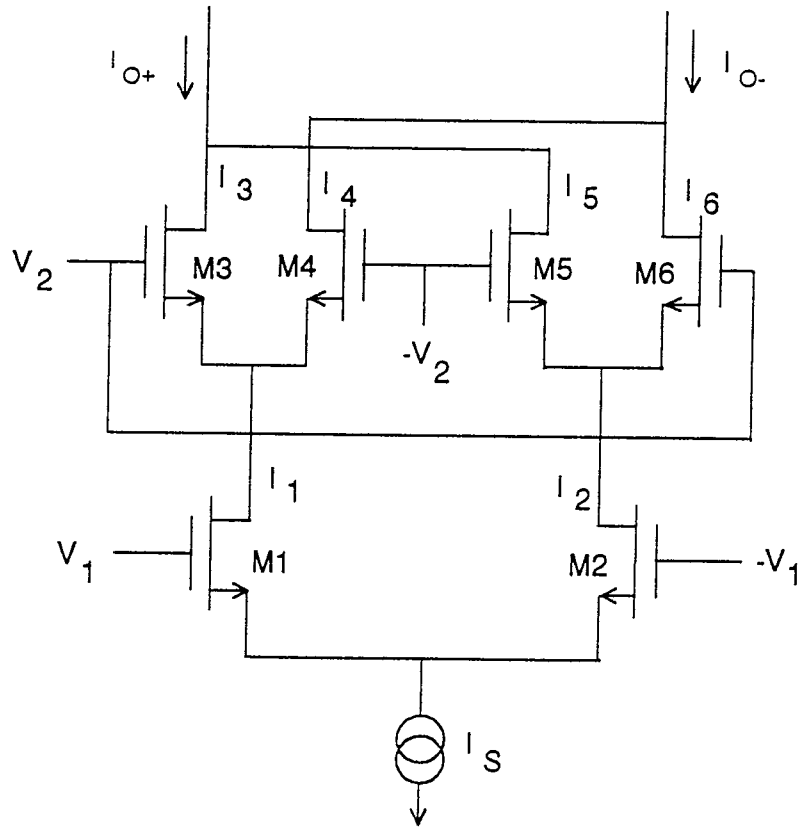


Figure 6.4: Standard Gilbert multiplier circuit.

The currents I_1 and I_2 satisfy

$$I_1 = K_1(V_1 + V_{01})^2 \tag{6.2}$$

$$I_2 = K_1(-V_1 + V_{01})^2 \tag{6.3}$$

$$I_S = I_1 + I_2 \tag{6.4}$$

where $V_{01} = \sqrt{\frac{I_S}{2K_1} - V_1^2}$ incorporates the shift in source voltage of M1 and M2 when the input V_1 is applied. Similarly, we can write equations for the currents I_3 through I_6:

$$I_3 = K_2(V_2 + V_{03})^2 \tag{6.5}$$

$$I_4 = K_2(-V_2 + V_{03})^2 \tag{6.6}$$

$$I_5 = K_2(-V_2 + V_{05})^2 \tag{6.7}$$

$$I_6 = K_2(V_2 + V_{05})^2 \tag{6.8}$$

$$I_1 = I_3 + I_4 \tag{6.9}$$

$$I_2 = I_5 + I_6 \tag{6.10}$$

where $V_{03} = \sqrt{\frac{I_1}{2K_2} - V_2^2}$ incorporates the shift in source voltage of M3 and M4 when the input V_2 is applied with I_1 flowing through M1, and $V_{05} = \sqrt{\frac{I_2}{2K_2} - V_2^2}$ incorporates the shift in source voltage of M5 and M6 when the input V_2 is applied with I_2 flowing through M2. Noting that

$$V_{03} = \sqrt{\frac{I_1}{2K_2} - V_2^2} \tag{6.11}$$

$$= \sqrt{\frac{K_1}{2K_2}(V_1 + V_{01})^2 - V_2^2} \tag{6.12}$$

$$V_{05} = \sqrt{\frac{I_2}{2K_2} - V_2^2} \tag{6.13}$$

$$= \sqrt{\frac{K_1}{2K_2}(-V_1 + V_{01})^2 - V_2^2} \tag{6.14}$$

we can rewrite the equations for the currents I_3 through I_6 as

$$I_3 = K_2\left(V_2+\sqrt{\frac{K_1}{2K_2}(V_1+V_{01})^2-V_2^2}\right)^2 \tag{6.15}$$

$$I_4 = K_2\left(-V_2+\sqrt{\frac{K_1}{2K_2}(V_1+V_{01})^2-V_2^2}\right)^2 \tag{6.16}$$

$$I_5 = K_2\left(-V_2+\sqrt{\frac{K_1}{2K_2}(-V_1+V_{01})^2-V_2^2}\right)^2 \tag{6.17}$$

$$I_6 = K_2\left(V_2+\sqrt{\frac{K_1}{2K_2}(-V_1+V_{01})^2-V_2^2}\right)^2. \tag{6.18}$$

We can now derive the output current.

$$I_O = I_{O+}-I_{O-} \tag{6.19}$$

$$= I_3-I_4+I_5-I_6 \tag{6.20}$$

$$\begin{aligned} = K_2\Bigg[&\left(V_2+\sqrt{\frac{K_1}{2K_2}(V_1+V_{01})^2-V_2^2}\right)^2 \\ &-\left(-V_2+\sqrt{\frac{K_1}{2K_2}(V_1+V_{01})^2-V_2^2}\right)^2 \\ &+\left(-V_2+\sqrt{\frac{K_1}{2K_2}(-V_1+V_{01})^2-V_2^2}\right)^2 \\ &-\left(V_2+\sqrt{\frac{K_1}{2K_2}(-V_1+V_{01})^2-V_2^2}\right)^2\Bigg] \end{aligned} \tag{6.21}$$

$$\begin{aligned} = K_2\Bigg[&2V_2\sqrt{\frac{K_1}{2K_2}(V_1+V_{01})^2-V_2^2} \\ &-2V_2\sqrt{\frac{K_1}{2K_2}(-V_1+V_{01})^2-V_2^2}\Bigg] \end{aligned} \tag{6.22}$$

$$\begin{aligned} = 2K_2V_2\Bigg[&\sqrt{\frac{K_1}{2K_2}}(V_1+V_{01})-\frac{1}{2}\sqrt{\frac{2K_2}{K_1}}\frac{V_2^2}{(V_1+V_{01})}+\dots \\ &-\sqrt{\frac{K_1}{2K_2}}(-V_1+V_{01})+\frac{1}{2}\sqrt{\frac{2K_2}{K_1}}\frac{V_2^2}{(-V_1+V_{01})}-\dots\Bigg] \end{aligned} \tag{6.23}$$

$$= 2K_2V_2\left[2\sqrt{\frac{K_1}{2K_2}}V_1 + \frac{1}{2}\sqrt{\frac{2K_2}{K_1}}V_2^2\left(\frac{1}{-V_1+V_{01}} - \frac{1}{V_1+V_{01}}\right)+\ldots\right] \quad (6.24)$$

$$= 2K_2V_2\left[\sqrt{\frac{2K_1}{K_2}}V_1 + \sqrt{\frac{2K_2}{K_1}}V_2^2\frac{V_1}{V_{01}^2-V_1^2}+\ldots\right] \quad (6.25)$$

$$\approx 2\sqrt{2K_1K_2}V_2V_1. \quad (6.26)$$

The final approximation is valid provided V_1 is small compared to V_{01}, $V_{01} > 1$, V_2 is small, and $\frac{K_2}{K_1}$ is modest. Thus, the circuit does indeed approximate a multiplier, with the output current as a function of input voltage given by $I_O = 2\sqrt{2K_1K_2}V_2V_1$.

Figure 6.5 shows how the wide-range Gilbert multiplier circuit compares to the standard Gilbert multiplier of figure 6.4. Transistors M1 through M6 serve as M1 through M6 of figure 6.4. Transistors M7 and M8 form a current mirror acting as the current source I_S of figure 6.4. M9 and M10 form a current mirror to drive the sources of M3 and M4 with I_1 from M1. Similarly, M11 and M12 form a current mirror to drive the sources of M5 and M6 with I_2 from M2.

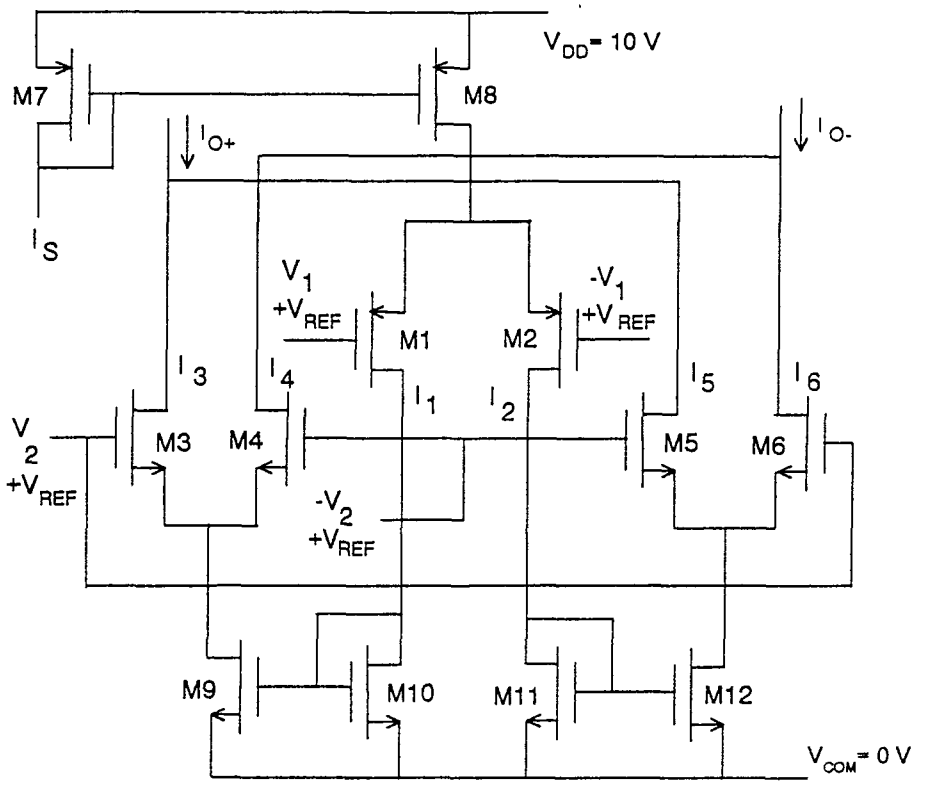


Figure 6.5: Wide-range Gilbert multiplier circuit for analysis.

Deviations from the ideal multiplier characteristic are caused by the terms neglected in the derivation of the multiplier characteristic, Early effect, nonideal characteristics of the current mirrors, and threshold mismatches (offsets) among the multiplier core transistors M1 through M6. Body effect will also cause deviations from the ideal characteristic, but can be reduced by putting as many of M1 through M6 as possible in their own wells. For

the circuit of figure 6.5 in a 2 μm MOSIS p-well technology, M3-M6 can be placed in separate p-wells.

6.3 Analog Low-Pass Filter Design

The analog building blocks needed for the implementation of the network of oscillators based on either the quadrature oscillator approach of figure 5.1 or the automatic gain control approach of figure 5.2 for the complex saturating nonlinearity include multipliers, summers, low-pass filters, and in the case of figure 5.1, a real saturating nonlinearity as well. Multiplier design has already been discussed, and summation of currents at multiplier outputs has been shown to not require any extra circuitry. (Voltages can be summed using a circuit even simpler than the multiplier circuit of figure 6.4, so even though summing currents is trivial, summing voltages is not difficult.) Real saturating nonlinearities for the quadrature oscillator approach can take the form of nonlinear resistive elements (such as MOS diodes or p-n diodes), or can consist of a small circuit. However, the low-pass filter design requires some consideration, because a low-pass filter can take up much more silicon real estate than a multiplier. Furthermore, there are constraints on what approaches can be used to achieve various time constants.

As a starting point for design of a low-pass filter in MOSIS 2 μm

double-polysilicon p-well technology, we will take the maximum size of a poly-poly capacitor (a subjective decision based on keeping the capacitor size on roughly the same scale as two multipliers) to be 3pF. Using a p-well resistor, resistances up to about 3MΩ can be achieved in a reasonable area for this application. Thus, a direct RC time constant down to about 20 kHz can be achieved.

Another approach for implementing the low-pass filter is to use the transconductance-C circuit of figure 6.6. The time constant for this circuit is approximately given by g_m/C where g_m is the small-signal transconductance of M3 and M4. An advantage of the transconductance-C circuit over the direct RC circuit is that the transconductance-C circuit can provide voltage gain. The circuit of figure 6.6 can be used to generate time constants down to the 10s of kHz range.

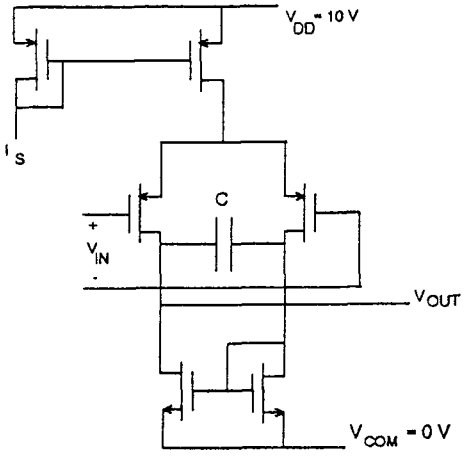


Figure 6.6: Transconductance-C low-pass filter.

The circuit of figure 6.7, known as a transconductance-killer circuit because the cross-coupled input stage acts to reduce the overall transconductance below that of the individual transistors, has been used to produce time constants down to 1 kHz. For lower corner frequencies, continuous-time approaches tend to require too much chip area, and switched-capacitor circuits should be considered.

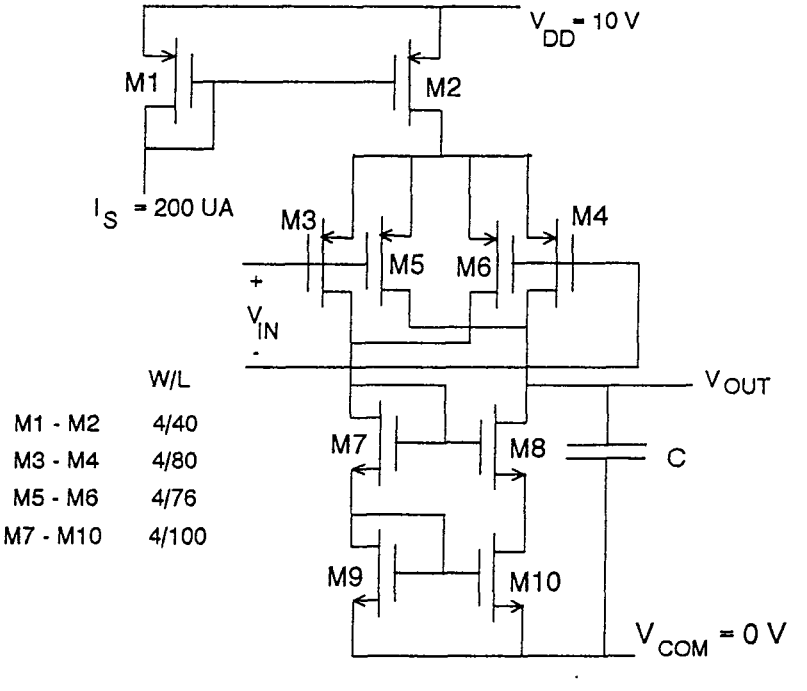


Figure 6.7: Transconductance-killer low-pass filter.

Of course, the low-pass filter design must be a simple pole. Any proposed low-pass filter design must be analyzed or simulated to make sure it will behave as a single-pole filter out to a high enough frequency. P-well resistors

have distributed capacitance, and the transconductance-C circuits of figures 6.6 and 6.7 have additional dynamics which must be taken into account. Obviously the low-pass filters just listed are by no means the only ones which could be considered. For example, an active op-amp low-pass filter can trade off increased power dissipation to produce a more ideal single-pole filter.

6.4 Tradeoffs among Power Dissipation, Network Low-Pass-Filter Corner Frequency, and Physical Circuit Size

The purpose of describing the multiplier circuit at length and of introducing some low-pass-filter circuits is that they enable tradeoffs among power dissipation, network low-pass-filter corner frequency, and physical circuit size to be explored. This is a first step toward determining potential applications for analog implementations of such networks.

To consider the tradeoff between power dissipation and network low-pass-filter corner frequency, some assumptions are needed. First, it will be assumed that the multipliers dominate the power dissipation. Second, the total power dissipation allowed will be taken to be 1 W (roughly the power level at which in a room temperature environment without heat sinking the package will feel warm to the touch, but not actually hot enough to burn a

finger). Finally, for the circuit of figure 6.1, with $I_s = 200\mu$A corresponding to a power dissipation of 4 mW per multiplier, the bandwidth at which "slight" deterioration of the frequency response is seen to occur is taken to be 20 MHz (multiplier circuits cannot be characterized by a single 3dB point, because the frequency response is different for the two inputs). This estimate is based on testing larger adaptive filter circuits of which the multiplier of figure 6.1 was a subcircuit.

As the multiplier current is decreased, its bandwidth will also decrease, but because reducing the current level requires changing the transistor sizes in order to maintain the same input signal range, the relationship between multiplier current and bandwidth can only be understood with the aid of the derivation of the multiplier characteristic given in section 6.2. The relationship between current and bandwidth can be estimated by taking bandwidth to be proportional to g_m, which depends both on transistor W/L and on drain current according to $g_m = 2\sqrt{KI_D} = \sqrt{2\mu C_{ox}\frac{W}{L}I_D}$. But in order for the approximation (6.26) to be valid, $V_{01} = \sqrt{\frac{I_S}{2K_1} - V_1^2}$ must remain constant as I_S is decreased; therefore, K_1 must decrease in proportion to I_S. But to keep the multiplier gain constant, K_2 must then also decrease in proportion to I_S. Hence the W/L of all the multiplier core transistors must be reduced in proportion to the reduction in I_S in order to maintain the same range

and gain. Therefore, since K is proportional to I_D and g_m is proportional to $\sqrt{KI_D}$, g_m is proportional to I_D. Thus, for the multiplier circuits of figures 6.4 and 6.5, bandwidth reduction is approximately linear with decreasing current.

Changing W/L will have an effect on transistor capacitances as well as on g_m, but the bandwidth changes due to decreasing I_D (using the 4mW multiplier as a starting point) will be primarily due to the change in g_m. However, as I_D is increased, there comes a point where due to Early effect or short-channel limitations L cannot be decreased any further so that increasing W/L means increasing only W. Overlap capacitances increase linearly with W, so increasing current increases both g_m and the transistor capacitances. Therefore, instead of bandwidth increasing linearly with I_D, it will increase only as the square root of I_D. Moreover, finite driving impedances and Miller effect act to prevent further increase in bandwidth with increasing current after a certain point. The 20MHz multiplier is probably near the point of diminishing returns so that increasing multiplier power above 4mW for the wide-range Gilbert configuration would not substantially improve its bandwidth performance.

If we assume a sparcely interconnected network of coupled oscillators, and assume four multipliers per unit are used to effect the complex saturating

nonlinearity while an average of four more (real) multipliers per unit are used for the complex interconnecting weights, then with a maximum power dissipation of 1 W, at 4mW per multiplier we can build roughly a 30-unit network.

It is more difficult to give a clear-cut tradeoff between number of multipliers and total silicon real estate than was the case for power dissipation. This is because low-pass filters tend to require more area than multipliers due to the capacitors, and interconnections can also consume a significant amount of area for a large number of units. For the adaptive filter project, two multipliers and two low-pass filters with 3pF capacitors (accompanied by switching and biasing circuitry) were integrated in a $250\mu m \times 1185\mu m$ cell size. Therefore, $1mm^2$ per unit would be a reasonable estimate of space required. A 30-unit network would therefore fit on a 6mm $\times$ 6mm chip, which can be accomodated by MOSIS.

Now that the tradeoffs between component bandwidth, power dissipation, and circuit size have been given some perspective, the relationship between component bandwidth and network low-pass-filter corner frequency is all that remains to describe. There are several factors which need to be taken into account, because how far the automatic gain control loop pole must be below the multiplier bandwidth depends on the gain of the automatic gain

control loop, and how far the network low-pass-filter corner frequency must be below the non-dominant network pole depends on the gain of the function $r(\cdot)$ near zero and on the weight values. A very rough estimate, for purposes of determining potential applications only, is that the non-dominant network pole (or the automatic gain control loop pole) should be down a decade from the frequency at which the multiplier begins to show "slight" frequency response deterioration, and the dominant pole (the network low-pass-filter corner frequency) should be down a decade from the non-dominant pole frequency. Thus, the 4mW multiplier would give roughly a 200kHz network. Of course, with good simulation tools and higher-speed processes than MOSIS, much higher bandwidths could ultimately be obtainable.

Chapter 7

Conclusions and Recommendations for Future Work

To summarize, convergence results for a class of networks of nonlinear coupled oscillators (or directional units, depending on the point of view taken) have been presented. Physical motivation has been given for the simplest network considered, but the convergence results for the rest of the networks are based solely on the adaptability of the original Lyapunov function arguments to the various networks. Also, a simple example circuit illustrates that such networks can be used in feedback configurations, with convergence properties that can be deduced simply if the circuit design follows certain rules. In particular, two distinct approaches for eliminating undesired stable equilibria due to unit amplitudes going to zero were presented. Next, analog VLSI implementation considerations were discussed for such networks. Tradeoffs among power dissipation, bandwidth, and network size were presented to assist in determining potential applications for analog VLSI networks of coupled oscillators.

One contribution of this work was to point out that the dynamics (1.1)-

(1.2) converge as a consequence of the convergence properties of the dynamics (2.12). One direction for future work would be to undertake a study of to what extent, if at all, other types of coupled oscillator networks can be made to conform to the convergence results presented here. In particular, one could attempt to extend the convergence results to the situation where the oscillators do not share a common frequency for all time.

Another direction for future work could consist of developing more general techniques (than Theorems 4.1 and 4.2) for determining the equilibrium points and whether they are stable. Then the dependence of the equilibria upon external weight controls could be specified so that the stable equilibria of the network could be programmed using control inputs. An example of a situation where this might be useful would be in smoothly transitioning between different gaits in a walking robot.

Finally, there is an interesting implication of the adaptively controlled network convergence result in terms of hierarchical control of coupled oscillator networks. If two or more networks are arranged in a hierarchical fashion, with the units of one network serving as the weights for the next network in the hierarchy, the convergence result implies that feedback can be applied from lower levels to higher levels and stability will be maintained, as long as the adaptive control law is obeyed. Consideration of networks of coupled os-

cillators for hierarchical control systems with local and global feedback might therefore be worthwhile.

Appendix A

Calculation: $\beta a_j r(\beta a_j) - \log(2\pi I_0(\beta a_j)) \to \infty$ as $a_j \to \infty$

To summarize the calculation, first, we let $m = \beta a_j$, and rewrite the expression ignoring the constant term $-\log 2\pi$:

$$mr(m) - \ln I_0(m) = f_1(m). \tag{A.1}$$

Clearly $\beta a_j r(\beta a_j) - \log(2\pi I_0(\beta a_j)) \to \infty$ as $a_j \to \infty$ if and only if $f_1(m) \to \infty$ as $m \to \infty$. Next, we split $f_1(m)$ into two terms, as follows:

$$\begin{aligned} f_1(m) &= mr(m) - \ln I_0(m) \\ &= (m - \ln I_0(m)) - m(1 - r(m)) \\ &= f_2(m) - f_3(m). \end{aligned} \tag{A.2}$$

Then we show that $f_2(m) \to \infty$ as $m \to \infty$ and $f_3(m) < 1 \ \ \forall m$. Then we may conclude that $f_1(m) \to \infty$ as $m \to \infty$.

Now we show that $f_2(m) = m - \ln I_0(m) = m \ln e - \ln I_0(m) = -\ln e^{-m} I_0(m) \to \infty$ as $m \to \infty$. The asymptotic behavior of $I_0(m)$ as $m \to \infty$ is given by [28]:

$$I_0(m) \sim \frac{e^m}{\sqrt{2\pi m}} \left\{ 1 + \frac{1}{8m} + \frac{9}{2!(8m)^2} + \cdots \right\}. \tag{A.3}$$

Thus, for m large enough, $e^{-m}I_0(m)$ behaves like $\frac{1}{\sqrt{2\pi m}}$.

We can proceed rigorously as follows: For M sufficiently large, we have for $m \geq M$,

$$\begin{aligned} e^{-m}I_0(m) &< \frac{2}{\sqrt{2\pi m}} \text{ (assume } M > 1) \\ \frac{1}{e^{-m}I_0(m)} &> \frac{\sqrt{2\pi m}}{2} = \sqrt{\frac{\pi m}{2}} \\ f_2(m) = \ln \frac{1}{e^{-m}I_0(m)} &> \ln \sqrt{\frac{\pi m}{2}} = \frac{1}{2} \ln \frac{\pi m}{2} \to \infty \text{ as } m \to \infty. \end{aligned} \quad \text{(A.4)}$$

Thus, $f_2(m) \to \infty$ as $m \to \infty$.

Next we show that $f_3(m) = m\,(1 - r(m)) < 1 \ \ \forall m$.

$$\begin{aligned} f_3(m) &= m\,(1 - r(m)) \\ &= m\left(1 - \frac{I_1(m)}{I_0(m)}\right) \\ \frac{df_3}{dm} &= \left(1 - \frac{I_1(m)}{I_0(m)}\right) + m\left(\frac{-\frac{d}{dm}\left(I_1(m)\right) I_0(m) + I_1^2(m)}{I_0^2(m)}\right) \\ &\quad \left(\text{note: } \tfrac{d}{dm}\left(mI_1(m)\right) = mI_0(m)\right) \\ &= 1 - \frac{I_1(m)}{I_0(m)} - \frac{m\frac{d}{dm}\left(I_1(m)\right)}{I_0(m)} + m\left(\frac{I_1(m)}{I_0(m)}\right)^2 \\ &= 1 - \frac{I_1(m)}{I_0(m)} - \frac{mI_0(m) - I_1(m)}{I_0(m)} + m\left(\frac{I_1(m)}{I_0(m)}\right)^2 \\ &= 1 - m + mr^2(m) \\ &= 1 - m\left(1 - r^2(m)\right). \end{aligned} \quad \text{(A.5)}$$

Observe that:

$$\begin{aligned} r(m) &< 1 \\ r^2(m) &< r(m) \\ 1 - r^2(m) &> 1 - r(m) \\ m\left(1 - r^2(m)\right) &> m\left(1 - r(m)\right) \\ 1 - \frac{df_3}{dm} &> f_3(m) \\ \frac{df_3}{dm} &< 1 - f_3(m). \end{aligned} \tag{A.6}$$

Let $f_4(m)$ be defined by:

$$\frac{df_4}{dm} = 1 - f_4(m) \tag{A.7}$$

$$f_4(0) = f_3(0) = 0.$$

Then by a simple integration, $f_4(m) = 1 - e^{-m}$. Since $f_3(0) = f_4(0)$, $f_3(m)$ is continuously differentiable, and $\frac{df_3}{dm} < \frac{df_4}{dm}$ $\forall m$, it follows that

$$f_3(m) < f_4(m) = 1 - e^{-m} < 1 \ \ \forall m. \tag{A.8}$$

Thus, $f_3(m) < 1$ $\forall m$.

By having shown $f_2(m) \to \infty$ as $m \to \infty$ and $f_3(m) < 1$ $\forall m$, we may now conclude that, indeed,

$$\beta a_j r(\beta a_j) - \log\left(2\pi I_0(\beta a_j)\right) \to \infty \ \ as \ \ a_j \to \infty. \tag{A.9}$$

Appendix B

Convergence property shared by fixed-weight networks

Two sets of dynamics for fixed-weight networks of coupled oscillators were presented. The first set was (2.21)-(2.23), repeated below:

$$\dot{y}_j = -y_j + r(\beta|x_j|)\frac{x_j}{|x_j|} \tag{B.1}$$

$$x_j = \sum_{k=1}^{n} y_k w_{jk}^* \tag{B.2}$$

$$\text{at equilibrium, } y_j = r(\beta|x_j|)\frac{x_j}{|x_j|}. \tag{B.3}$$

The second set was (2.24)-(2.26), repeated below:

$$\dot{x}_j = -x_j + \sum_k y_k w_{jk}^* \tag{B.4}$$

$$y_j = r(\beta|x_j|)\frac{x_j}{|x_j|} \tag{B.5}$$

$$\text{at equilibrium, } x_j = \sum_{k=1}^{n} y_k w_{jk}^* \tag{B.6}$$

(this is the network of Zemel et. al.)

The convergence proof of section 2.3 applies directly to the second set of dynamics; thus, the x_j and y_j variables in the second set of dynamics converge to an equilibrium point. As shown by equation (2.3), the x_j variables in the first set of dynamics satisfy the same differential equation as the x_j

variables in the second set of dynamics. Therefore, all that needs to be proved to conclude that the first set of dynamics converges to an equilibrium point is that $x_j(t) \to x_j^*$ as $t \to \infty$ with $\dot{y}_j = -y_j + r(\beta|x_j|)\frac{x_j}{|x_j|}$ implies $y_j(t) \to y_j^* = r(\beta|x_j^*|)\frac{x_j^*}{|x_j^*|}$ as $t \to \infty$. This will follow from the following lemma:

Lemma B.1: Let $\frac{du}{dt} = -u + f(t)$, where $\lim_{t\to\infty} f(t) = 0$. Then $\lim_{t\to\infty} u(t) = 0$.

Proof: Observe that if $|f(t)| < \epsilon \ \ \forall t \geq T$, then for $t \geq T$,

$$u(t) = e^{-(t-T)}u(T) + \int_T^t e^{-(t-\tau)}f(\tau)d\tau \tag{B.7}$$

$$|u(t)| \leq e^{-(t-T)}|u(T)| + \epsilon\int_T^t e^{-(t-\tau)}d\tau \tag{B.8}$$

$$= e^{-(t-T)}|u(T)| + \epsilon(1 - e^{-(t-T)}) \tag{B.9}$$

$$\leq e^{-(t-T)}|u(T)| + \epsilon. \tag{B.10}$$

So given ϵ_1, we can find T_1 such that $|f(t)| < \frac{\epsilon_1}{2} \ \ \forall t \geq T_1$. For any initial condition $u(0)$, we can compute

$$u(T_1) = e^{-T_1}u(0) + \int_0^{T_1} e^{-(T_1-\tau)}f(\tau)d\tau \tag{B.11}$$

and then take $T_2 \geq T_1$ large enough that $e^{-(T_2-T_1)}|u(T_1)| < \frac{\epsilon_1}{2}$. Then $\forall t > T_2$, $|u(t)| < \epsilon_1$. Hence, $\lim_{t\to\infty} u(t) = 0$. Q.E.D.

To apply this lemma, let $u(t) = y_j(t) - r(\beta|x_j^*|)\frac{x_j^*}{|x_j^*|}$ and $f(t) = r(\beta|x_j|)\frac{x_j}{|x_j|} - r(\beta|x_j^*|)\frac{x_j^*}{|x_j^*|}$. Then $u(t)$ satisfies $\dot{u}(t) = -u(t) + f(t)$, and $f(t) \to 0$ as $t \to \infty$. Hence, $u(t) \to 0$ as $t \to \infty$, and we may conclude that $y_j(t) \to r(\beta|x_j^*|)\frac{x_j^*}{|x_j^*|}$ as $t \to \infty$.

Appendix C

Derivation of the Implementable Version of the Adaptively Controlled Network

Let

$$\dot{y}_j = -y_j + r(\beta|x_j|)\frac{x_j}{|x_j|} \tag{C.1}$$

$$x_j = \sum_{k=1}^{n} y_k w^*_{jk} + U_j \tag{C.2}$$

$$I^j_{l_1,l_2} = \begin{cases} 1 & : \quad w_{l_1,l_2} = d_{l_1,l_2} r(\beta|x_j|)\frac{x_j}{|x_j|} \\ 0 & : \quad \text{otherwise.} \end{cases} \tag{C.3}$$

Then

$$\begin{aligned} \dot{x}_j &= \sum_k \dot{y}_k w^*_{jk} + \sum_k y_k \dot{w}^*_{jk} + \dot{U}_j \\ &= \sum_k \dot{y}_k w^*_{jk} + \sum_{l_1,l_2} I^{l_1}_{j,l_2} y_{l_2} \dot{w}^*_{j,l_2} + \dot{U}_j \\ &= \sum_k \left(-y_k + r(\beta|x_k|)\frac{x_k}{|x_k|} \right) w^*_{jk} + \sum_{l_1,l_2} I^{l_1}_{j,l_2} y_{l_2} \dot{w}^*_{j,l_2} + \dot{U}_j \\ &= -x_j + \sum_k r(\beta|x_k|)\frac{x_k}{|x_k|} w^*_{jk} + \sum_{l_1,l_2} I^{l_1}_{j,l_2} y_{l_2} \dot{w}^*_{j,l_2} + \dot{U}_j + U_j. \end{aligned} \tag{C.4}$$

If we define the dynamics of U_j as

$$\dot{U}_j = -U_j - \sum_{l_1,l_2} I^{l_1}_{j,l_2} y_{l_2} \dot{w}^*_{j,l_2} + \sum_{l_1,l_2} I^{j}_{l_1,l_2} d_{l_1,l_2} r(\beta|x_{l_1}|) r(\beta|x_{l_2}|) \frac{x_{l_2}}{|x_{l_2}|} \frac{x^*_{l_1}}{|x_{l_1}|} \tag{C.5}$$

we obtain

$$\dot{x}_j = -x_j + \sum_k r(\beta|x_k|)\frac{x_k}{|x_k|}w^*_{jk} + \sum_{l_1,l_2} I^j_{l_1,l_2} d_{l_1,l_2} r(\beta|x_{l_1}|)r(\beta|x_{l_2}|)\frac{x_{l_2}}{|x_{l_2}|}\frac{x^*_{l_1}}{|x_{l_1}|}. \tag{C.6}$$

Because we wish to rewrite the dynamics in a form homogeneous in the x variables, let

$$W_j = U_j + \sum_{l_1,l_2} I^{l_1}_{j,l_2} y_{l_2} w^*_{j,l_2}. \tag{C.7}$$

Then

$$\begin{aligned}
\dot{W}_j &= \dot{U}_j + \sum_{l_1,l_2} I^{l_1}_{j,l_2}\dot{y}_{l_2}w^*_{j,l_2} + \sum_{l_1,l_2} I^{l_1}_{j,l_2} y_{l_2}\dot{w}^*_{j,l_2} \\
&= -U_j + \sum_{l_1,l_2} I^j_{l_1,l_2} d_{l_1,l_2} r(\beta|x_{l_1}|)r(\beta|x_{l_2}|)\frac{x_{l_2}}{|x_{l_2}|}\frac{x^*_{l_1}}{|x_{l_1}|} + \sum_{l_1,l_2} I^{l_1}_{j,l_2}\dot{y}_{l_2}w^*_{j,l_2} \\
&= -W_j + \sum_{l_1,l_2} I^j_{l_1,l_2} d_{l_1,l_2} r(\beta|x_{l_1}|)r(\beta|x_{l_2}|)\frac{x_{l_2}}{|x_{l_2}|}\frac{x^*_{l_1}}{|x_{l_1}|} \\
&\quad + \sum_{l_1,l_2} I^{l_1}_{j,l_2}(\dot{y}_{l_2} + y_{l_2})w^*_{j,l_2} \\
&= -W_j + \sum_{l_1,l_2} I^j_{l_1,l_2} d_{l_1,l_2} r(\beta|x_{l_1}|)r(\beta|x_{l_2}|)\frac{x_{l_2}}{|x_{l_2}|}\frac{x^*_{l_1}}{|x_{l_1}|} \\
&\quad + \sum_{l_1,l_2} I^{l_1}_{j,l_2} r(\beta|x_{l_2}|)\frac{x_{l_2}}{|x_{l_2}|}w^*_{j,l_2} \\
&= -W_j + \sum_{l_1,l_2} \left(I^j_{l_1,l_2} d_{l_1,l_2} + I^{l_1}_{j,l_2} d_{j,l_2}\right) r(\beta|x_{l_1}|)r(\beta|x_{l_2}|)\frac{x_{l_2}}{|x_{l_2}|}\frac{x^*_{l_1}}{|x_{l_1}|}.
\end{aligned} \tag{C.8}$$

Thus, the implementable dynamics are

$$\dot{y}_j = -y_j + r(\beta|x_j|)\frac{x_j}{|x_j|} \tag{C.9}$$

$$x_j = \sum_{k=1}^{n} y_k w^*_{jk} + W_j - \sum_{l_1,l_2} I^{l_1}_{j,l_2} d_{j,l_2} y_{l_2} r(\beta|x_{l_1}|)\frac{x^*_{l_1}}{|x_{l_1}|} \tag{C.10}$$

$$\dot{W}_j = -W_j + \sum_{l_1,l_2} \left(I^j_{l_1,l_2} d_{l_1,l_2} + I^{l_1}_{j,l_2} d_{j,l_2} \right) r(\beta|x_{l_1}|) r(\beta|x_{l_2}|) \frac{x_{l_2}}{|x_{l_2}|} \frac{x^*_{l_1}}{|x_{l_1}|} \tag{C.11}$$

at equilibrium, $y_j = r(\beta|x_j|)\frac{x_j}{|x_j|}$. (C.12)

References

[1] Avis H. Cohen and P. Wallen. The Neuronal Correlate of Locomotion in Fish. *Experimental Brain Research*, 41:11-18, 1980.

[2] Avis H. Cohen, ed. *Neural Control of Rhythmic Movements In Vertibrates.* John Wiley & Sons, Inc., 1988.

[3] Avis H. Cohen, Philip J. Holmes, and Richard H. Rand. The Nature of the Coupling Between Segmental Oscillators of the Lamprey Spinal Generator for Locomotion: A Mathematical Model. *Journal of Mathematical Biology*, 13:345-369, 1982.

[4] Nancy Kopell and G. B. Ermentrout. Coupled Oscillators and the Design of Central Pattern Generators. *Mathematical Biosciences*, 90:87-109, 1988.

[5] Daryl Marvin. Coordination of the Walking Stick Insect Using a System of Nonlinear Coupled Oscillators. Institute for Systems Research M.S. Thesis Report, M.S. 92-9, 1992.

[6] Sanjay Mishra. Control of Biped Locomotion Using Oscillators. Institute for Systems Research M.S. Thesis Report, M.S. 93-12, 1993.

[7] R. Eckhorn, W. Bauer, M. Brosch, W. Kruse, M. Munk, and H.J. Reitboeck. Coherent Oscillations: A Mechanism of Feature Linking in the Visual Cortex? *Biological Cybernetics*, 60:121-130, 1988.

[8] Charles Gray and Wolf Singer. Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proc. National Academy of Science, U.S.A.*, 86:1698-1702, March 1989.

[9] Pierre Baldi and Ronny Meir. Computing with Arrays of Coupled Oscillators: An Application to Preattentive Texture Discrimination. *Neural Computation*, 2:458-471, 1990.

[10] Erik Lumer and Bernardo Huberman. Binding Hierarchies: A Basis for Dynamic Perceptual Grouping. *Neural Computation*, 4:341-355, 1992.

[11] Michael Mozer, Richard Zemel, Marlene Behrmann, and Christopher Williams. Learning to Segment Images Using Dynamic Feature Binding. *Neural Computation*, 4:650-665, 1992.

[12] Lokendra Shastri and Venkat Ajjanagadde. From simple associations to systematic reasoning: A connectionist representation of rules, variables, and dynamic bindings using temporal synchrony. *Behavioral and Brain Sciences*, 1992(?).

[13] Roger Brocket. Pattern Generation and Feedback Control of Nonholonomic Systems. Preprint, 1993.

[14] Naomi Leonard. Averaging and Motion Control on Lie Groups. Institute for Systems Research Technical Research Report 93-59, 1993.

[15] Richard Zemel, Christopher Williams, and Michael Mozer. Lending Direction to Neural Networks. Preprint, 1993(?).

[16] K.V. Mardia. *Statistics of Directional Data.* New York: Academic Press, 1972.

[17] Hassan Khalil. *Nonlinear Systems.* New York: Macmillan Publishing Co., 1992.

[18] Emanuel Fischer. *Intermediate Real Analysis.* New York: Springer-Verlag, 1983.

[19] Solomon Lefschetz. *Differential Equations: Geometric Theory 2nd Ed.* New York: Wiley Interscience, 1963.

[20] Randall Geiger, Phillip Allen, and Noel Strader. *VLSI Design Techniques for Analog and Digital Circuits.* New York: McGraw Hill, 1990.

[21] Fritz Kub and Eric Justh. Analog CMOS Implementation of High Frequency Least Mean Square Error Learning Circuits. To appear in the IEEE International Solid-State Circuits Conference, Feb. 1995.

[22] Fritz Kub and Eric Justh. High Frequency Analog Circuits Implementing Tapped Delay Lines and the LMS Algorithm. *Proc. Long Island Adaptive Antenna Systems Symposium,* pp. 121-123, 1994.

[23] Fritz Kub, Eric Justh, Francis Long, and Keith Moon. High-Frequency Adaptive Learning Element. *Government Microcircuits Applications Conference (GOMAC),* pp. 291-294, 1992.

[24] Fritz Kub, Keith Moon, and Eric Justh. Analog VLSI Neural Network Integrated Circuits. *Technology 2001 Conference*, pp. 147-155, 1991.

[25] Phillip Allen and Douglas Holberg. *CMOS Analog Circuit Design.* New York: Holt, Rinehart, and Winston, 1987.

[26] Paul Horowitz and Winfield Hill. *The Art of Electronics, 2nd ed.* New York: Cambridge University Press, 1989.

[27] Allen Lytel. *Handbook of Algebraic & Trigonometric Functions.* Indianapolis: Howard W. Sams & Co., 1964.

[28] W.G. Bickley. *Bessel Functions and Formulae.* Cambridge: University Press, 1957.