Biol. Cybernetics 22, 73—84 (1976)

Asymptotics Applied to a Neural Network

Jack W. Silverstein

Division of Applied Mathematics and Center for Neural Studies, Brown University, Providence, Rhode Island, USA

Received: July 10, 1975

Abstract

A mathematical model of neural processing is proposed which incorporates a theory for the storage of information. The model consists of a network of neurons that linearly processes incoming neural activity. The network stores the input by modifying the synaptic properties of all of its neurons. The model lends support to a distributive theory of memory using synaptic modification. The dynamics of the processing and storage are represented by a discrete system. Asymptotic analysis is applied to the system to show the learning capabilities of the network under constant input. Results are also given to predict the network's ability to learn periodic input, and input subjected to small random fluctuations.

1. Introduction

1.1. Basic Considerations

Neural processing is loosely defined as the conversion of neural activity of one group of neurons into the activity of another group. The network presented in this paper models mechanisms underlying neural processing. It brings together into mathematical form prominent physiological and psychological ideas of learning and memory. These ideas can be stated under the following general areas:

1) Neural Plasticity. Memory storage is believed to be the result of physical changes in neurons. The question of neural plasticity has been extensively investigated, especially at the synaptic level [1–4], and memory models have been developed which assume synaptic modification [5, 6, 8–10].

2) Distributive Memory. There is evidence that a large number of neurons is used in storing a memory trace [11], where an individual neuron plays a small role, but is involved in storing many traces. This *distributive* theory of memory is similar to properties of holograms and has led to several holographic type models [12, 13].

3) Memory Consolidation. It is believed that the permanent storage of an experience does not occur immediately, but is dependent on the persistence, or "reverberation", of neural activity initiated by the experience. This is referred to as the *consolidation hypothesis* of memory [14]. The reverberating activity, thought to occur in neural "circuits", is associated with the notion of *short term memory.*

The network proposed will be assumed to process neural activity linearly. Considering the threshold properties of the action potential this seems to be a major simplification. However, there is evidence that linearity is a good approximation to the processing of spiking frequencies of neurons encoding sensory input [15]. We take the activity of a neuron to be the *deviation* of spiking frequency from the spontaneous frequency the neuron exhibits at rest. In this way positive and negative activity can be associated with excitation and inhibition respectively.

1.2. The Model

The assumptions and ideas above are incorporated into the following model of neural processing:

Let $\mathscr{I}$, $\mathscr{N}$, and $\mathscr{O}$ represent three mutually disjoint sets of neurons. The neurons in $\mathscr{I}$ synapse onto some of the neurons in $\mathscr{N}$, and some neurons in $\mathscr{N}$ synapse onto the neurons in $\mathscr{O}$. The activity in $\mathscr{I}$ is therefore processed onto $\mathscr{O}$ via $\mathscr{N}$, and $\mathscr{N}$ is defined as the network. We can imagine the location of $\mathscr{I}$ to be either peripheral to the brain or somewhere in the brain. In the first case $\mathscr{I}$, for example, could be the ganglion cells making up the optic nerve that preprocesses visual stimuli. In the second case $\mathscr{I}$ could be cells in cortex projecting information to another cortical area.

Let $\mathscr{I}$ and $\mathscr{O}$ contain n neurons each, where n is large. For a given time $t \in \{0, 1, 2, \ldots\}$, let the *input vector* $y(t) \in R^n$ denote the activity in $\mathscr{I}$. More specific, if $\mathscr{I} = \{\mathscr{I}_1, \ldots, \mathscr{I}_n\}$, then at time t the activity of $\mathscr{I}_i$ is $y_i(t)$. Similarly, let the *output vector* $x(t) \in R^n$ denote the activity in $\mathscr{O}$ at time t. We assume the cells in $\mathscr{N}$ undergo changes in time, and since the processing it performs is done linearly, we associate a linear operator $M(t): R^n \to R^n$ with $\mathscr{N}$ at time t. The elements in $\mathscr{O}$ are *fed back* into $\mathscr{N}$ in such a manner as to have:

$$x(t+1) = M(t)(y(t) + \alpha x(t)), \tag{1.2.1}$$

where $\alpha \in R$ is called the amplification factor (see (1.2.2)).

$$\begin{tikzcd} \mathscr{I} \arrow[r] & \mathscr{N} \arrow[r] & \mathscr{O} \arrow[l, bend left] \end{tikzcd} \quad . \tag{1.2.2}$$

As we will see, short term memory can be associated with $x(t)$. Long term effects are achieved by the way $M(t)$ is altered. We investigate a modification scheme in which the occurence of each $y(t)$ is stored so that $\mathcal{N}$ is able to remember which inputs it has "seen". The scheme involves a function h mapping R^n into the set of $n \times n$ matrices. The matrix $M(t)$ changes according to the equation:

$$M(t+1) = \varrho M(t) + \sigma h(y(t) + \alpha x(t))$$
$$0 \leqq \varrho, \sigma \leqq 1 \qquad \varrho + \sigma = 1 \tag{1.2.3}$$

where ϱ is called the forgetting coefficient, and σ is the modification coefficient.

In Section 2 we introduce the *Synapse Modification Model*, originally formulated by Anderson [5, 6]. The matrix $M(t)$ stores information by forming projection operators of subspaces spanned by the inputs. We will see how this modification scheme is related to coordinated changes in synaptic properties of $\mathcal{N}$. It is a distributive model in the sense that information is stored among many synapses.

Our goal is to understand mathematically the recognition properties of the *Synapse Modification Model*. This paper is the first step toward the task. The next two sections contain some essential mathematical concepts and tools used in Section 2 to analyze the network's learning capabilities. It will be shown that the network is capable of learning a constant input. Evidence will also be given to support the conjecture that the network can learn periodic input, as well as input subjected to small random fluctuations.

Another modification scheme has been proposed in which synaptic properties are kept fixed but transmission properties of cells are able to change in a subset of $\mathcal{N}$ (this scheme for learning has been suggested by Pfaff [16]). The structure of the network in this case is in a form in which we are able to investigate properties of randomly connected networks [17]. At $t=0$ connections are generated according to a controlled probability model. In that paper it is shown that, under reasonable assumptions, the spectral measure of $M(0)$ converges in probability to a universal one as n tends to infinity.

1.3. Discrete Systems and Stability

From $t=0$ onward, $(x(t), M(t))$, with $(x(0), M(0)) = (x_0, M_0)$ is a solution to a discrete system, also called an ordinary difference equation; that is, it is a solution to an equation of the form:

$$z(t+1) = F(t, z(t)) \tag{1.3.1}$$

where $z: \{0, 1, 2, \ldots\} \to \mathcal{B}$, $F: \{0, 1, 2, \ldots\} \times \mathcal{B} \to \mathcal{B}$ is continuous, and $\mathcal{B}$ is a finite dimensional Euclidean space. We write $z(t, t_0, z_0)$ as the solution to (1.3.1) with $z(t_0, t_0, z_0) = z_0$. For each t, $(x(t), M(t))$ can be considered to be an $n^2 + n$-dimensional vector, so that $\mathcal{B} = R^{n^2+n}$. It is clear that F in (1.2.1), (1.2.3) depends on h and the sequence of inputs $Y \equiv \{y(t)\}_{t=0}^{\infty}$.

For non-linear systems such as (1.2.1), (1.2.3), it is unlikely that solutions can be found in closed form. One of the few theories we can apply to (1.3.1) concerns the asymptotic behavior of solutions. We use the theory on the *Synapse Modification Model* for special cases of Y. Section 2.4 deals in limiting properties of (1.2.1), (1.2.3), when $Y = \{y\}_{t=0}^{\infty}$, that is, $y(t) = y$, $t = 0, 1, 2, \ldots$. In this case F is independent of t and we have:

$$z(t+1) = F(z(t)) \,. \tag{1.3.2}$$

This is called an autonomous difference equation. We write $z(t, z_0)$ as the unique solution of (1.3.2) with $z(0, z_0) = z_0$. Each solution can be viewed as a collection of points in $\mathcal{B}$, called an orbit.

Let $\| \; \|$ be the Euclidean norm on $\mathcal{B}$, that is, if $x = (x_1, \ldots, x_n)$ then $\|x\| = \left(\sum_{i=1}^{n} x_i^2 \right)^{1/2}$. We will use the following definitions:

Definition. Let $F(b) = b$. The discrete system (1.3.2) is *stable* at b if $\forall \varepsilon > 0 \, \exists \delta > 0$ such that $\|z_0 - b\| \leqq \delta \Rightarrow \|z(t, z_0) - b\| \leqq \varepsilon \, \forall t \geqq 0$.

In words this means the system is stable at b if for any neighborhood N_1 of b, there is another neighborhood N_2 of b in which any solution to the system starting in N_2 will eventually be trapped in N_1.

Definition. The system (1.3.2) is *asymptotically stable* at b if it is stable at b and there is a neighborhood N of b such that $z_0 \in N \Rightarrow z(t, z_0) \to b$ as $t \to \infty$.

Definition. The system (1.3.2) is unstable at b if it is not stable there. The point b, where $F(b) = b$, is called a *critical point* of (1.3.2).

We will make use of the following theorem which can be seen in Hahn [18]:

Theorem 1.3. Suppose $F(0) = 0$ and is differentiable at 0. Let $dF(0) = \left(\frac{\partial f_i}{\partial z_j} (z=0) \right)$ be the differential of F at 0. Then

1. If all the eigenvalues of $dF(0)$ are all less than one in absolute value, then (1.3.2) is asymptotically stable at 0.
2. If at least one eigenvalue of $dF(0)$ is greater than one in absolute value, then (1.3.2) is unstable at 0.

The theorem can be used for any b such that $F(b) = b$, since we simply let $v(t) = z(t) - b$, so that

$$v(t+1) = F(v(t) + b) - b \,. \tag{1.3.3}$$

1.4. Existence of Bounded and Periodic Solutions

In Section 2.5 we apply the following theorem which is a special case of a theorem in Hurt [19]:

Theorem 1.4. Suppose A is an $n \times n$ matrix with eigenvalues less than one in absolute value. Let $\| \ \|$ be a norm on R^n where the sup norm it induces on $n \times n$ matrices $\left(\text{that is, } B\ n \times n \text{ matrix} \Rightarrow \|B\|_{\sup} = \sup_{\|x\|=1} \|Bx\|\right)$ is such that $\|A\|_{\sup} < 1$.

Let $F: \{\ldots, -1, 0, 1, \ldots\} \times R^n \to R^n$ satisfy:

1. For some $N > 0$, $\|F(t, z)\| \leqq N(1 - \|A\|_{\sup})$ for all $t \in \{\ldots, -1, 0, 1, \ldots\}$ and all z such that $\|z\| \leqq N$.
2. There exists $F_1 < 1 - \|A\|_{\sup}$ such that for all z_1, z_2 with $\|z_i\| \leqq N$, $i = 1, 2$, and for all $t \in \{\ldots -1, 0, 1, \ldots\}$ we have:

$$\|F(t, z_1) - F(t, z_2)\| \leqq F_1 \|z_1 - z_2\| . \tag{1.4.1}$$

Then there exists a bounded solution $z^*(t)$ of:

$$z(t+1) = Az(t) + F(t, z(t)) \qquad t \in \{\ldots, -1, 0, 1, \ldots\} \tag{1.4}$$

where $\|z^*(t)\| \leqq N$ for all t. This is the only bounded solution with $\|z^*(t)\| \leqq N\, \forall t$. Moreover, if $F(t, z)$ is periodic in t with period T, then $z^*(t)$ is periodic with period T.

We will see in Section 2.5 that conditions can be put on the inputs to ensure bounded (or periodic) solutions. Since we are looking at solutions from $t=0$ onwards, the uniqueness property will not hold for bounded F, but it will hold for F periodic in t since there is only one way to extend F backwards in time so that it will be periodic for all $t \in \{\ldots -1, 0, 1, \ldots\}$.

2. Synapse Modification Model

2.1. Derivation

Let $\mathcal{N}$ consist of two disjoint sets of neurons $\mathcal{N}^1 = \{\mathcal{N}_1^1, \ldots \mathcal{N}_n^1\}$ and $\mathcal{N}^2 = \{\mathcal{N}_1^2, \ldots, \mathcal{N}_n^2\}$. We assume each $\mathcal{N}_j^1 \in \mathcal{N}^1$ makes synaptic contact with every $\mathcal{N}_i^2 \in \mathcal{N}^2$, that is, a synapse exists between a dendritic branch of $\mathcal{N}_i^2$ and an axonal branch of $\mathcal{N}_j^1$. Therefore, neural activity is being processed from $\mathcal{N}^1$ to $\mathcal{N}^2$. Let $\mathcal{I} = \{\mathcal{I}_1, \ldots, \mathcal{I}_n\}$ and $\mathcal{O} = \{\mathcal{O}_1, \ldots, \mathcal{O}_n\}$. We assume that for each $i \in \{1, \ldots, n\}$, $\mathcal{N}_i^2$ and only $\mathcal{N}_i^2$ makes synaptic contact with $\mathcal{O}_i$. Also, for $j \in \{1, \ldots, n\}$, $\mathcal{I}_j$ and $\mathcal{O}_j$ both make synaptic contact with $\mathcal{N}_j^1$, and these are the only neurons synapsing onto $\mathcal{N}_j^1$. Figure 1 shows in schematic form the wiring of $\mathcal{I}$, $\mathcal{N}$, and $\mathcal{O}$, for $n=3$.

At time $t \in \{0, 1, \ldots\}$ we associate a number $m_{ij}(t)$ with the synapse between $\mathcal{N}_j^1 \in \mathcal{N}^1$ and $\mathcal{N}_i^2 \in \mathcal{N}^2$, called the *synaptic strength* between these cells. Let $b_j^1(t)$ and $b_i^2(t)$ be the activities at time t of $\mathcal{N}_j^1$ and $\mathcal{N}_i^2$ respectively. Processing of $\mathcal{I}$ and $\mathcal{O}$ onto $\mathcal{N}^1$ and of $\mathcal{N}^2$ onto $\mathcal{O}$ is assumed to occur with no time delay,

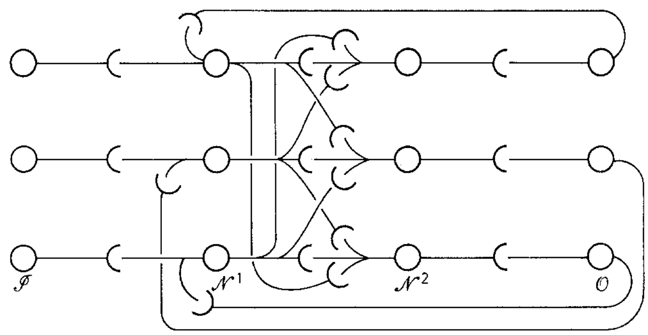

Fig. 1. Schematic diagram of the network for $n=3$

and is to consist only of translating activity. Recalling the definitions of the *input vector* $y(t)$ and the *output vector* $x(t)$ we have:

$$b_j^1(t) = y_j(t) + \alpha x_j(t), \qquad j = 1, \ldots, n \tag{2.1.1}$$

$$x_i(t) = b_i^2(t), \qquad i = 1, \ldots, n. \tag{2.1.2}$$

Processing of $\mathcal{N}^1$ onto $\mathcal{N}^2$ is assumed to be governed by the following equation:

$$b_i^2(t+1) = \sum_{j=1}^{n} m_{ij}(t) b_j^1(t), \qquad i = 1, \ldots, n. \tag{2.1.3}$$

Let $M(t) = (m_{ij}(t))$. Using (2.1.1) and (2.1.2), and putting (2.1.3) in vector notation we get (1.2.1).

We assume $m_{ij}(t)$ changes according to

$$\begin{aligned} m_{ij}(t+1) &= \varrho m_{ij}(t) + \sigma b_i^1(t) b_j^1(t), \quad 0 \leqq \varrho, \sigma \leqq 1, \quad \varrho + \sigma = 1 \\ &= \varrho m_{ij}(t) + \sigma(y_i(t) + \alpha x_i(t))(y_j(t) + \alpha x_j(t)). \end{aligned} \tag{2.1.4}$$

Therefore we arrive at the equations governing the *Synapse Modification Model*:

$$\begin{cases} x(t+1) = M(y(t) + \alpha x(t)) \\ M(t+1) = \varrho M(t) + \sigma (y(t) + \alpha x(t))(y(t) + \alpha x(t))^T \\ 0 \leqq \varrho, \sigma \leqq 1 \qquad \varrho + \sigma = 1 \qquad \alpha \in R. \end{cases} \tag{2.1.5}$$

The symbol T stands for the transpose operation.

2.2. Physical Significance

This model assumes that processing is being performed by taking a weighted sum of activity in $\mathcal{I}$, where the weights are the synaptic strengths associated with the synapses. There seems to be no evidence for or against synapses participating in neural processing in this way.

We have no evidence supporting the modification scheme either. Synaptic strength could be related to some aspect of a synapse known to be involved in the efficacy of transmission. For example, the amount of transmitter released by the presynaptic cell, the influence the transmitter substance has on depolarizing the post-synaptic membrane, or the amount of pre- and post-synaptic contact. The activities of $\mathcal{N}_i^1$ and $\mathcal{N}_j^1$ are responsible for altering the synaptic strengths between: $\mathcal{N}_i^1$ and $\mathcal{N}_j^2$; $\mathcal{N}_i^1$ and $\mathcal{N}_i^2$; $\mathcal{N}_j^1$ and $\mathcal{N}_i^2$; $\mathcal{N}_j^1$ and $\mathcal{N}_j^2$. Is this physically realizable? We offer one possible explanation:

Suppose: 1. Modification of some aspect of a synapse depends on the simultaneous presence of a quantity of substance S_1 on the presynaptic and of a quantity of another substance S_2 on the post-synaptic side.

2. The quantity of S_1 depends on the activity of the presynaptic cell.

3. All $\mathcal{N}_i^1 - \mathcal{N}_i^2$ connections are special in the sense that the quantity of S_2 on all synaptic sights of the dendrite of $\mathcal{N}_i^2$ depends on the activity of $\mathcal{N}_i^1$.

Then the appropriate modifications could be achieved.

The third assumption replaces the backwards information hypothesis needed to support Anderson's association model [5, 6] which has been studied by Cooper [8], and Grossberg's model [9, 10]. In Anderson's model the pairing of an input vector f with an output vector g is stored in memory by adding the outer product gf^T to the linear operator. The synapse connecting cells $\mathcal{N}_i^1$ and $\mathcal{N}_j^2$ alters due to the activity in $\mathcal{N}_i^1$ and $\mathcal{N}_j^2$. Some way of receiving knowledge of the activity in $\mathcal{N}_j^2$ must be assumed. Assumption three seems more feasible than the backwards information hypothesis.

2.3. A Model for Recognition

The main reason for investigating this modification scheme (2.1.4) can be seen from the following:

Suppose $M(0)=0$, inputs $y(t)$ come into $\mathcal{N}$ at each $t\in\{0, 1, \dots, T\}$ and no inputs come in after T for some time, that is, $y(t)=0$ for $t>T$ and less than some T_1. Suppose $x(T_1)$ is negligible due to the amplification factor α ($|\alpha|<1$, say). At time T_1 a vector y is processed. If it is orthogonal to all previous inputs, then $x(T_1+1)=0$. If not, then $\|x(T_1+1)\|$ is a non-zero quantity, the size depending mainly on the relation between y and the previous inputs. We see then that $\mathcal{N}$ will *recognize* previous inputs or linear combinations of them. If T_1 is much larger than T then, due to the forgetting coefficient ϱ, $\|x(T_1+1)\|$ will be small, indicating the vague remembrance of y.

Due to the feedback of $x(t)$ an input y seen for a few moments will linger around for quite some time afterwards. For example, if y appears at time t, then at time $t+1$, σyy^T is added to $\varrho M(t)$. Then $x(t+2)$ will contain $\sigma yy^T(y(t+1)+\alpha x(t))=\sigma<y, y(t+1)+\alpha x(t)>y$. Another multiple of $P_y\equiv yy^T\div\|y\|^2$ (called the *projection operator* of the space spanned by y) is added to $\varrho M(t+1)$, and so on. It is clear then that $x(t)$ can be associated with a short term memory mechanism.

The assumption restricting α to the real domain and $\varrho+\sigma=1$ are imposed mainly for mathematical convenience. It seems more reasonable to replace α with a matrix. However, there should be some relationship between ϱ and σ. We simply assume that the amount forgotten is the same as the amount stored. We feel ϱ should be close to 1, that is, little forgetting should occur over a short period of time. One part of the analysis in Section 2.4 depends on the size of ϱ. Too many cases would have to be considered if ϱ is 2/3 or less. So we shall assume that ϱ is greater than 2/3.

The next three sections contain the mathematical analysis done thus far on the *Synapse Modification Model*. In Section 2.4 we investigate the stability properties of (2.1.5) under constant input. In Section 2.5 we derive some conditions for the existence of bounded and periodic solutions. Section 2.6 looks at the model subjected to random input. Section 2.7 contains some concluding remarks.

2.4. Constant Input

We begin with some definitions and obvious lemmas involving outer products.

Definition. $\langle\ \rangle$ denotes the Euclidean inner product in R^n, i.e. $x, y\in R^n\Rightarrow\langle x, y\rangle=\sum_{i=1}^{n} x_iy_i$.

Definition. $\|\ \|$ denotes the Euclidean norm on R^n, i.e. $\|x\|=(\langle x, x\rangle)^{1/2}$.

Definition. If A is an $n\times n$ matrix, then $\|A\|_{\sup}$ is the sup norm induced by the Euclidean norm on R^n, i.e. $\|A\|_{\sup}=\sup_{\|x\|=1}\|Ax\|$.

Definition. $|A|$ denotes the determinant of A.

Lemma 1. $x, y, z\in R^n\Rightarrow(xy^T)z=x(y^Tz)=\langle y, z\rangle x$.

Lemma 2. $x\in R^n\Rightarrow xx^T$ is a symmetric matrix in R^n with eigenvalues $\|x\|^2$ and 0. $\|x\|^2$ has multiplicity one and 0 has multiplicity $n-1$.

Lemma 3. $x, y\in R^n\Rightarrow\|xy^T\|_{\sup}=\|x\|\,\|y\|$.

Let $y(t)=y$ $t\in\{0, 1, \dots\}$. Then (2.1.5) becomes:

$$\begin{cases} x(t+1)=M(t)(y+\alpha x(t)) \\ M(t+1)=\varrho M(t)+\sigma(y+\alpha x(t))(y+\alpha x(t))^T \\ 0\leqq\varrho, \sigma\leqq 1, \quad \varrho+\sigma=1 \quad \alpha\in R. \end{cases} \tag{2.4.1}$$

Let $x(0)=x_0$, $M(0)=M_0$.

We assume $\varrho < 1$[1]. If $\alpha = 0$, then (2.4.1) becomes:

$$\begin{cases} x(t+1) = M(t)y \\ M(t+1) = \varrho M(t) + \sigma yy^T \end{cases} \tag{2.4.2}$$

so that

$$\begin{cases} M(t) = \varrho^t M_0 + \sigma \left(\sum_{s=0}^{t-1} \varrho^s \right) yy^T & t > 1 \\ x(t) = \varrho^t M_0 y + \sigma \|y\|^2 \left(\sum_{s=0}^{t-1} \varrho^s \right) y . \end{cases} \tag{2.4.3}$$

Therefore:

$$M(t) \to \sigma \left(\sum_{s=0}^{\infty} \varrho^s \right) yy^T = \frac{\sigma}{1-\varrho} yy^T = yy^T \quad \text{as} \quad t \to \infty \tag{2.4.4}$$

and

$$x(t) \to \|y\|^2 y \quad \text{as} \quad t \to \infty . \tag{2.4.5}$$

For $\alpha \neq 0$, we calculate the critical points of (2.4.1). Assume $y \neq 0$. The case $y = 0$ will be dealt with shortly. We solve:

$$\begin{cases} x = M(y + \alpha x) \\ M = \varrho M + \sigma (y + \alpha x)(y + \alpha x)^T \end{cases} \tag{2.4.6}$$

or

$$\begin{cases} x = M(y + \alpha x) \\ M = (y + \alpha x)(y + \alpha x)^T . \end{cases} \tag{2.4.7}$$

Therefore:

$$x = \|y + \alpha x\|^2 (y + \alpha x) \tag{2.4.8}$$

which gives us

$$(1 - \alpha \|y + \alpha x\|^2) x = \|y + \alpha x\|^2 y . \tag{2.4.9}$$

$(1 - \alpha \|y + \alpha x\|^2)$ cannot be zero, since if it was $\|y + \alpha x\|^2$ would be zero and that would imply $1 = 0$. Also $\|y + \alpha x\|^2$ cannot be zero, since if it was x would be $-\frac{1}{\alpha} y$ and 0 at the same time.

Therefore, x is a non-zero multiple of y, say $x = x_0 y$. Then M must be a multiple of yy^T, say $M = m_0 yy^T$.

From the first equation in (2.4.7) we get

$$x_0 y = m_0 yy^T (y + \alpha x_0 y) = m_0 \|y\|^2 y + \alpha x_0 m_0 \|y\|^2 y \tag{2.4.10}$$

so that

$$x_0 (1 - \alpha m_0 \|y\|^2) = m_0 \|y\|^2 . \tag{2.4.11}$$

[1] We will not carry out the analysis of the case $\varrho = 1$, $\sigma = 0$, which corresponds to no learning. This case essentially results in a linear equation in $x(t)$. The limiting properties of solutions to the equation will depend on y and (x_0, M_0).

If $(1 - m_0 \|y\|^2)$ is zero, m_0 would be zero, which implies $1 = 0$. Therefore $(1 - m_0 \|y\|^2) \neq 0$ and we get

$$x_0 = \frac{m_0 \|y\|^2}{1 - \alpha m_0 \|y\|^2} . \tag{2.4.12}$$

From the second equation in (2.4.15) we get

$$m_0 yy^T = (1 + \alpha x_0)^2 yy^T \tag{2.4.13}$$

so that

$$m_0 = (1 + \alpha x_0)^2 = \left(1 + \alpha \left(\frac{m_0 \|y\|^2}{1 - \alpha m_0 \|y\|^2} \right) \right)^2$$
$$= \frac{1}{(1 - \alpha m_0 \|y\|^2)^2} > 0 . \tag{2.4.14}$$

Therefore all critical points are of the form:

$$\begin{cases} x = x_0 y \\ M = m_0 yy^T \end{cases} \tag{2.4.15}$$

where

$$m_0 (1 - \alpha m_0 \|y\|^2)^2 = 1$$
$$x_0 = \frac{m_0 \|y\|^2}{1 - \alpha m_0 \|y\|^2} . \tag{2.4.16}$$

Let $m = \|M\| = m_0 \|y\|^2$ (Lemma 3). Then we get:

$$m(1 - \alpha m)^2 = \|y\|^2 . \tag{2.4.17}$$

The graph of $f(m) = m(1 - \alpha m)^2$ for $\alpha > 0$ is shown in Fig. 2.

Since $f\left(\frac{1}{3\alpha}\right) = \frac{4}{27\alpha}$, it is clear from the graph that the following theorem holds:

Theorem 2.4.1. For $\alpha > 0$, there are three cases:

1. If $\|y\| > \frac{2}{3} \frac{1}{\sqrt{3\alpha}}$, then there is one critical point (x^1, M^1) where $\|M\| > \frac{1}{\alpha}$.

2. If $\|y\| = \frac{2}{3} \frac{1}{\sqrt{3\alpha}}$, then there are two critical points (x_1^2, M_1^2), (x_2^2, M_2^2) where $\|M_1^2\| = \frac{1}{3\alpha}$ and $\|M_2^2\| > \frac{1}{\alpha}$.

3. If $0 < \|y\| < \frac{2}{3} \frac{1}{\sqrt{3\alpha}}$, then there are three critical points (x_1^3, M_1^3), (x_2^3, M_2^3), (x_3^3, M_3^3), where $0 < \|M_1^3\| < \frac{1}{3\alpha} < \|M_2^3\| < \frac{1}{\alpha} < \|M_3^3\|$.

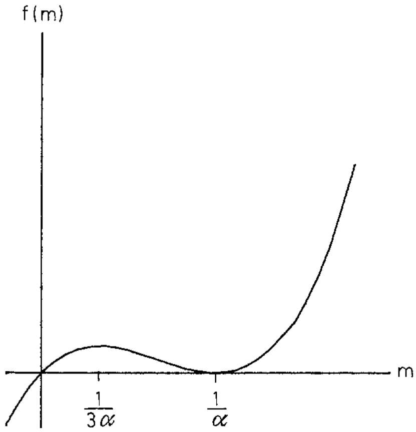

Fig. 2. Graph of $f(m)=m(1-\alpha m)^2$ for $\alpha>0$

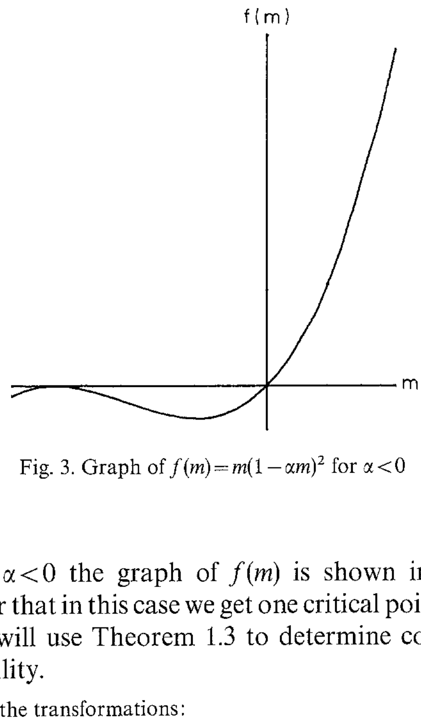

Fig. 3. Graph of $f(m)=m(1-\alpha m)^2$ for $\alpha<0$

For $\alpha<0$ the graph of $f(m)$ is shown in Fig. 3. It is clear that in this case we get one critical point (x, M).

We will use Theorem 1.3 to determine conditions for stability.

We make the transformations:

$$\begin{cases} x'(t)=x(t)-x \\ M'(t)=M(t)-M \end{cases} \tag{2.4.18}$$

where (x, M) is a critical point of (2.4.1). We consider then:

$$\begin{cases} x'(t+1)=\alpha M x'(t)+M'(t)(y+\alpha x)+\alpha M'(t)x'(t) \\ M'(t+1)=\varrho M'(t)+\sigma\alpha(y+\alpha x)x'(t)^T+\sigma\alpha x'(t)(y+\alpha x)^T+\sigma\alpha^2(t)x'(t)^T . \end{cases} \tag{2.4.19}$$

This can be converted into a system in n^2+n dimensional space by letting $M_i'(t)$ be the i^{th} column of $M'(t)$ and considering:

$$z(t)=\begin{bmatrix} x'(t) \\ M_1'(t) \\ \vdots \\ M_n'(t) \end{bmatrix} . \tag{2.4.20}$$

Let $w=y+\alpha x$. The differential D at $z=0$ of this huge system is:

$$D=\begin{bmatrix} \alpha w w^T & w_1 I & w_2 I \dots & w_n I \\ D_1 & \varrho I & 0 \dots & 0 \\ D_2 & 0 & \varrho I & 0 \\ \vdots & \vdots & \vdots & \vdots \\ D_n & 0 & 0 & \varrho I \end{bmatrix}$$

where I is the $n\times n$ identity matrix, the zeros denote the $n\times n$ 0 matrix, and:

$$D_i=\sigma\alpha\left(w_i I+\begin{bmatrix} 0\dots 0 & w_1 & 0\dots 0 \\ 0\dots 0 & w_2 & 0\dots 0 \\ \vdots & \vdots & \vdots \\ 0\dots 0 & w_n & 0\dots 0 \end{bmatrix}\right). \tag{2.4.22}$$

$\uparrow$
i^{th} column

The eigenvalues λ of D satisfy $p(\lambda)\equiv|D-\lambda I_{n^2+n}|=0$ where $p(\lambda)$ is the characteristic polynomial of D and I_{n^2+n} is the $n^2+n\times n^2+n$ identity matrix.

Assume for the moment $\lambda\neq\varrho$. Multiplying the $n+1, \dots, 2n$ rows of $D-\lambda I_{n^2+n}$ by $-w_1\div(\varrho-\lambda)$ and adding these rows to the first n rows will result in a matrix with zeros in the first n rows and $n+1, \dots, 2n$ columns. We continue in this manner until all elements in the first n rows and $n+1, \dots, n^2+n$ columns are zero. Then the determinant of the resulting matrix is just $(\varrho-\lambda)^{n^2}$ times the determinant of the $n\times n$ matrix in the upper left corner. This $n\times n$ matrix turns out to be $(\lambda\neq\varrho)$:

$$(\varrho-\lambda)^{-n}\left((2\varrho-1-\lambda)ww^T-\left((\varrho-\lambda)\frac{\lambda}{\alpha}+\sigma\|w\|^2\right)I\right). \tag{2.4.23}$$

Since $ww^T=M$ and $\|w\|^2=\|M\|$ we can write for $\lambda\neq\varrho$:

$$p(\lambda)=\alpha^n(\varrho-\lambda)^{n^2-n}\left|(2\varrho-1-\lambda)M-\left((\varrho-\lambda)\frac{\lambda}{\alpha}+\sigma\|M\|\right)I\right| . \tag{2.4.24}$$

Since $p(\lambda)$ is continuous in λ, (2.4.32) is true for $\lambda=\varrho$, which gives us n^2-n roots of $p(\lambda)$, all less than 1 in absolute value.

From Lemma 2, $(2\varrho-1-\lambda)M$ has 0 as an eigenvalue with multiplicity $n-1$, and has $(2\varrho-1-\lambda)\|M\|$ as an eigenvalue with multiplicity 1. Therefore, we get $2(n-1)$ roots of $p(\lambda)$ from the two roots of:

$$p_1(\lambda)\equiv\lambda^2-\varrho\lambda-\sigma\alpha\|M\| \tag{2.4.25}$$

and two roots of $p(\lambda)$ from the two roots of:

$$\begin{aligned} p_2(\lambda)&=\lambda^2-\varrho\lambda-\sigma\alpha\|M\|+\alpha(2\varrho-1-\lambda)\|M\| \\ &=\lambda^2-(\varrho+\alpha\|M\|)\lambda+\alpha(3\varrho-2)\|M\| . \end{aligned} \tag{2.4.26}$$

We have found all the roots of $p(\lambda)$ since $n(n-1)+2(n-1)+2=n^2+n$.

We now determine how the roots of $p_1(\lambda)$ and $p_2(\lambda)$ are in relation to the unit circle. Consider first $p_1(\lambda)$. If the roots of $p_1(\lambda)$ are real, if $p_1(-1)$ and $p_1(1)$ are positive, and if the minimum of $p_1(\lambda)$ lies between -1 and 1, then the roots of $p_1(\lambda)$ are less than one in absolute value. If $p_1(-1)$ or $p_1(1)$ is negative, then the roots of $p_1(\lambda)$ are real and at least one of them is greater than one in absolute value.

We will work through the case for $\alpha>0$. The minimum of $p_1(\lambda)$ occurs at $\varrho\div 2$ and its value there is:

$$p_1\left(\frac{\varrho}{2}\right)=-\frac{\varrho^2}{4}-\sigma\alpha\|M\|<0 . \tag{2.4.27}$$

Therefore, the roots of $p_1(\lambda)$ are real and the minimum lies between -1 and 1. We have

$$p_1(-1)=1+\varrho-\sigma\alpha\|M\|\gtrless 0 \Leftrightarrow \|M\|\lessgtr \frac{1+\varrho}{\alpha\sigma} \tag{2.4.28}$$

$$p_1(1)=1-\varrho-\sigma\alpha\|M\|\gtrless 0 \Leftrightarrow \|M\|\lessgtr \frac{1}{\alpha}. \tag{2.4.29}$$

We approach $p_2(\lambda)$ in the same way. The minimum of $p_2(\lambda)$ occurs at $(\varrho+\alpha\|M\|)\div 2$ and its value there is:

$$p_2\left(\frac{\varrho+\alpha\|M\|}{2}\right)=\frac{-(\varrho-\alpha\|M\|)^2}{4}-2\alpha\sigma\|M\|<0. \tag{2.4.30}$$

Therefore the roots of $p_2(\lambda)$ are real. We have:

$$p_2(-1)=1+\varrho+\alpha(3\varrho-1)\|M\|>0 \tag{2.4.31}$$

since we are assuming ϱ is greater than $\frac{2}{3}$ (Section 2.3), and

$$p_2(1)=1-\varrho-\alpha\|M\|+\alpha(3\varrho-2)\|M\|\gtrless 0 \Leftrightarrow \|M\|\lessgtr \frac{1}{3\alpha}. \tag{2.4.32}$$

When $p_2(1)>0$, $(\varrho+\alpha\|M\|)\div 2$ is between -1 and 1.

Comparing (2.4.28), (2.4.29), and (2.4.32), and using Theorem 1.3 we have:

Theorem 2.4.2. For $\alpha>0$, (2.4.1) is asymptotically stable at (x, M) if $\|M\|<\frac{1}{3\alpha}$. (2.4.1) is unstable at (x, M) if $\|M\|>\frac{1}{3\alpha}$.

Combining with Theorem 2.4.1 we have:

Theorem 2.4.3. For $\alpha>0$:

1. If $\|y\|>\frac{2}{3}\frac{1}{\sqrt{3\alpha}}$, the one critical point (x^1, M^1) is unstable.
2. If $\|y\|=\frac{2}{3}\frac{1}{\sqrt{3\alpha}}$, (x_2^2, M_2^2) is unstable.
3. If $0<\|y\|<\frac{2}{3}\frac{1}{\sqrt{3\alpha}}$, then (x_1^3, M_1^3) is asymptotically stable and (x_2^3, M_2^3) and (x_3^3, M_3^3) are both unstable.

For $\alpha<0$, using the assumption $\varrho>\frac{2}{3}$, we arrive at the following theorem:

Theorem 2.4.4. For $\alpha<0$ if $\|M\|<\frac{1+\varrho}{|\alpha|(3\varrho-1)}$, then the critical point (x, M) is asymptotically stable. If $\|M\|>\frac{1+\varrho}{|\alpha|(3\varrho-1)}$, then (x, M) is unstable.

We now consider $y=0$. We see from (2.4.7) that the critical points must satisfy:

$$\begin{cases} M=\alpha^2 xx^T \\ x=\alpha^3\|x\|^2 x \end{cases} \tag{2.4.33}$$

so that either $x=0$, $M=0$, or $\|x\|=|\alpha|^{-3/2}$ which in this case implies $\|M\|=\frac{1}{|\alpha|}$. For $\alpha>0$ we conclude that $x=0$, $M=0$ is asymptotically stable, and all other critical points are unstable.

For $\alpha<0$, since $\frac{1+\varrho}{3\varrho-1}<1$ we conclude that $x=0$, $M=0$ is asymptotically stable, and all other critical points are unstable.

Theorem 1.3 tells us nothing about the critical point (x_1^2, M_1^2) in case 2 of Theorem 2.4.1 ($\alpha>0$) or when $\|M\|=\frac{\varrho+1}{|\alpha|(3\varrho-1)}$ $(\alpha<0)$. From looking at phase portraits when $n=1$ we are quite certain this point is unstable. We have no way of analytically proving it at the moment.

It is possible to show, when $\alpha>0$, that the critical points form a region in which solutions are bounded. Equation (2.4.8) can be rewritten as:

$$x_0=(1+\alpha x_0)^3\|y\|^2. \tag{2.4.34}$$

Figure 4 contains the graph of $v=(1+\alpha u)^3$. A line going through the origin with slope $\frac{1}{\|y\|^2}$ will intersect this curve in one, two, or three places, depending on the size of $\|y\|^2$. Since the roots of (2.4.42) are a continuous function of $\|y\|$, we can conclude:

1. In case 1 of Theorem 2.4.1 $x^1=x_0 y$ with $x_0<0$.
2. In case 2 of Theorem 2.4.1 $x_1^2=x_{01}y$, $x_2^2=x_{02}y$ with $x_{01}>0$, $x_{02}<0$.
3. In case 3 of Theorem 2.4.1 $x_1^3=x_{01}y$, $x_2^3=x_{02}y$, $x_3^3=x_{03}y$ with $0<x_{01}<x_{02}$, $x_{03}<0$.

Suppose $(\boldsymbol{x}, \boldsymbol{M})=(x_0 y, m_0 yy^T)$ is a critical point with $x_0>0$, and let (x, M) be such that $\|x\|\leqq\|\boldsymbol{x}\|$, $\|M\|_{\sup}\leqq\|\boldsymbol{M}\|_{\sup}$. Then:

$$\begin{aligned}\|M(y+\alpha x)\| &\leqq \|M\|_{\sup}(\|y\|+\alpha\|x\|)\\ &\leqq \|\boldsymbol{M}\|_{\sup}(\|y\|+\alpha\|\boldsymbol{x}\|)\\ &= m_0\|y\|^2(\|y\|+\alpha x_0\|y\|)\\ &= (\text{from (2.4.10)})\quad \|\boldsymbol{x}\|\end{aligned} \tag{2.4.35}$$

and

$$\begin{aligned}\|\varrho M+\sigma(y+\alpha x)(y+\alpha x)^T\| &\leqq \varrho\|M\|_{\sup}+\sigma\|y+\alpha x\|^2\\ &\leqq \varrho\|\boldsymbol{M}\|_{\sup}+\sigma(\|y\|+\alpha\|\boldsymbol{x}\|)^2\\ &= \varrho m_0\|y\|^2+\sigma(1+\alpha x_0)^2\|y\|^2\\ &= (\text{from (2.4.14)})\\ &\quad (\varrho m_0+\sigma m_0)\|y\|^2=\|\boldsymbol{M}\|_{\sup}.\end{aligned} \tag{2.4.36}$$

Therefore, in case 2 of Theorem 2.4.1 all solutions starting in

$$B_2\equiv\{(x, M);\ \|x\|\leqq\|x_1^2\|,\ \|M\|_{\sup}\leqq\|M_1^2\|_{\sup}\} \tag{2.4.37}$$

stay in B_2, and in case 3 of Theorem 2.4.1 all solutions starting in

$$B_3 \equiv \{(x, M);\ \|x\| \leqq \|x_2^3\|,\ \|M\|_{\sup} \leqq \|M_2^3\|_{\sup}\} \tag{2.4.38}$$

stay in B_3.

We are quite sure these solutions will approach critical points. From Hurt [20] we conclude that cluster points of a bounded solution forms an invariant set S, that is, solutions starting in S will remain in S. If we could show that S contains only one point then we are done. We have not yet found a way of proving this.

For $n=1$, with $x, m, y \in R$, (2.4.1) becomes:

$$\begin{cases} x(t+1) = m(t)(y+\alpha x(t)) \\ m(t+1) = \varrho m(t) + \sigma(y+\alpha x(t))^2 . \end{cases} \tag{2.4.39}$$

Figure 5 is a phase portrait of (2.4.39) with $\alpha = .2$, $y = .5$, $\varrho = .9$.

We originally investigated the structure of phase portraits of (2.4.39) by looking at the analogous set of differential equations:

$$\begin{cases} \dot{x}(t) = m(t)(y+\alpha x(t)) - x(t) \\ \dot{m}(t) = (\varrho - 1)m(t) + \sigma(y+\alpha x(t))^2 . \end{cases} \tag{2.4.40}$$

The analogue of Theorem 1.3 for differential equations has been known for quite some time. The analysis of (2.4.40) also predicts the global structure of the phase portrait of (2.4.39) above the x axis. Intuitively, it seems that whenever the steps are small in (2.4.39) $x(t+1) - x(t)$ and $m(t+1) - m(t)$ are close to $\dot{x}(t)$ and $\dot{m}(t)$ in (2.4.40) so (2.4.39) and (2.4.40) should be similar.

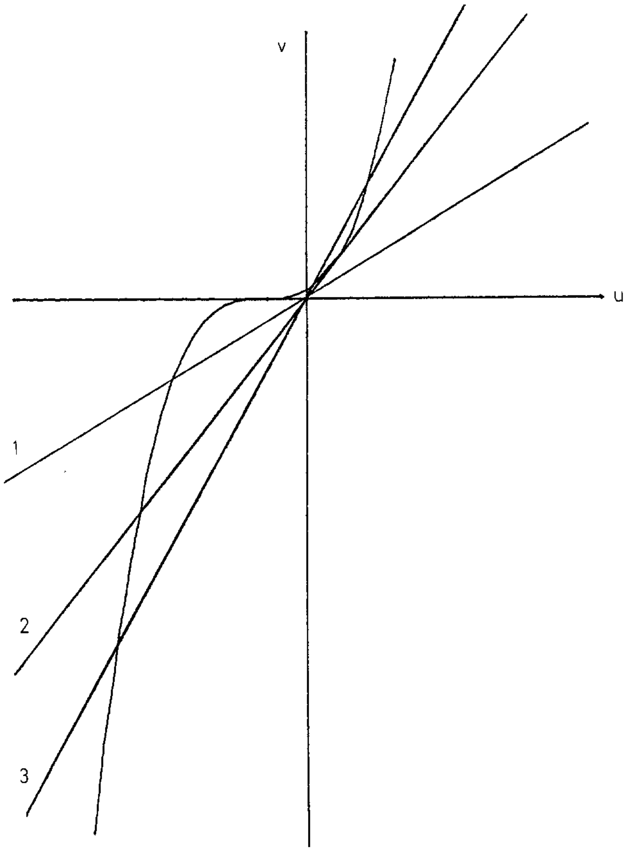

Fig. 4. Three lines intersecting the curve $v = (1+\alpha u)^3$. Each line is given by: $v = \frac{1}{\|y\|^2} u$, where 1. $\|y\| > \frac{2}{3} \frac{1}{\sqrt{3\alpha}}$, 2. $\|y\| = \frac{2}{3} \frac{1}{\sqrt{3\alpha}}$, 3. $\|y\| < \frac{2}{3} \frac{1}{\sqrt{3\alpha}}$

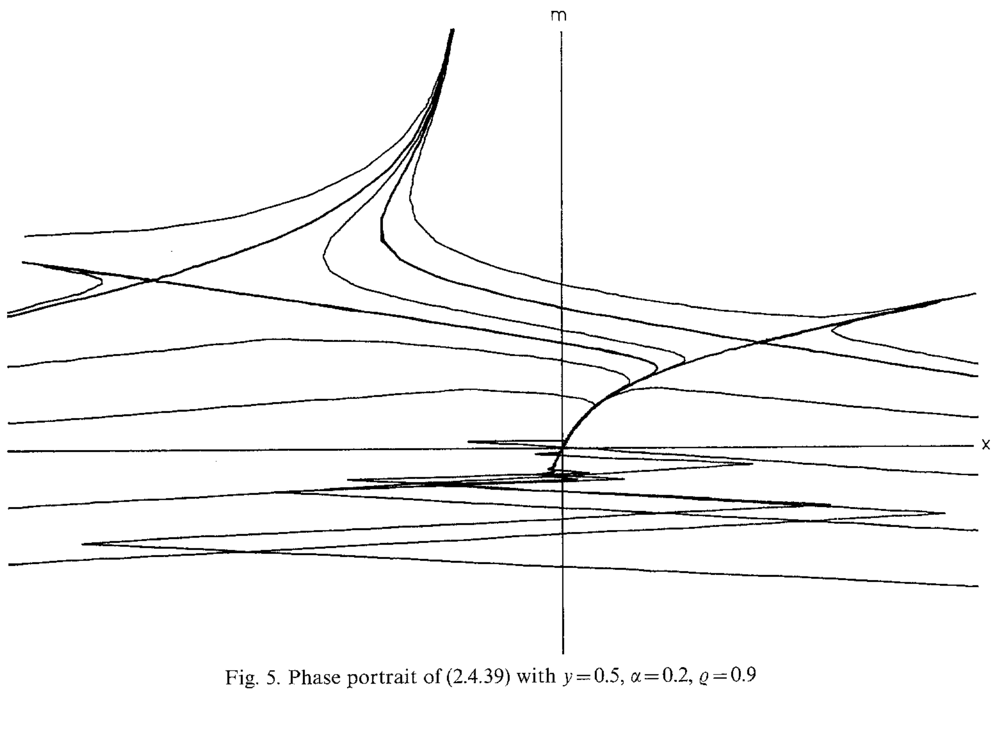

Fig. 5. Phase portrait of (2.4.39) with $y = 0.5$, $\alpha = 0.2$, $\varrho = 0.9$

Notice, though, the chaos occurring below the x axis. The steps are large due to the changing of sign of $x(t)$.

We do not know of any theory proving the global relationships of (2.4.39) and (2.4.40) but feel much more can be done with autonomous difference equations.

2.5. Bounded and Periodic Solutions

Suppose the inputs $y(t)$ $(t \geqq 0)$ are bounded or periodic with period T. We will use Theorem 1.4 to find conditions on the inputs and the parameters to ensure the existence of a bounded solution, and, if the inputs are T-periodic, this solution is periodic.

We will, as before, consider (x, M) to be a vector in R^{n^2+n} space (see 2.4.20) and we use the Euclidean norm. Let $\|M\|_E = \left(\sum_{ij} m_{ij}^2\right)^{1/2}$. It can be shown that $\|M\|_{\sup} \leqq \|M\|_E$. Then we have:

$$\|(x, M)\| = (\|x\|^2 + \|M\|_E^2)^{1/2}. \tag{2.5.1}$$

Notice that:

$$\|xy^T\|_E = \left(\sum_{ij} (x_i y_j)^2\right)^{1/2} = \left(\sum_i x_i^2 \sum_j y_j^2\right)^{1/2} = \|x\| \; \|y\|. \tag{2.5.2}$$

We let

$$F(t, (x, M)) = (M(y(t) + \alpha x), \sigma(y(t) + \alpha x)(y(t) + \alpha x)^T) \tag{2.5.3}$$

so that the matrix A is an $n^2 + n \times n^2 + n$ diagonal matrix:

$$A = \left[\begin{array}{c|cccc} 0 & & 0 & & \\ \hline & \varrho & 0 & 0 \ldots & 0 \\ & 0 & \varrho & 0 \ldots & 0 \\ 0 & \vdots & \vdots & \vdots & \vdots \\ & 0 & 0 & \ldots & \varrho \end{array}\right] \begin{array}{l} \} \; n \text{ rows} \\ \\ \left.\vphantom{\begin{array}{c}a\\a\\a\\a\end{array}}\right\} n^2 \text{ rows}. \end{array} \tag{2.5.4}$$

We see that $\|A\|_{\sup} = \varrho < 1$. Also

$$\begin{aligned} \|F(t, (x, M))\|^2 &= \|M(y(t) + \alpha x)\|^2 + \sigma^2 \|(y(t) + \alpha x)(y(t) + \alpha x)\|_E^2 \\ &= \|M(y(t) + \alpha x)\|^2 + \sigma \|y(t) + \alpha x\|^4 \\ &\leqq \|M\|_{\sup}^2 \|y(t) + \alpha x\|^2 + \sigma^2 \|y(t) + \alpha x\|^4 \\ &\leqq (\|M\|_E^2 + \sigma^2 \|y(t) + \alpha x\|^2) \|y(t) + \alpha x\|^2. \end{aligned} \tag{2.5.5}$$

For (x_1, M_1), (x_2, M_2) we have

$$\begin{aligned} &\|F(t, (x_1, M_1)) - F(t, (x_2, M_2))\|^2 \\ &= \|M_1(y(t) + \alpha x_1) - M_2(y(t) + \alpha x_2)\|^2 \\ &\quad + \sigma^2 \|(y(t) + \alpha x_1)(y(t) + \alpha x_1)^T - (y(t) + \alpha x_2)(y(t) + \alpha x_2)^T\|_E^2 \\ &= \|(M_1 - M_2)(y(t) + \alpha x_2) + \alpha M_1(x_1 - x_2)\|^2 \\ &\quad + \sigma^2 \|\alpha(y(t) + \alpha x_1)(x_1 - x_2)^T + \alpha(x_1 - x_2)(y(t) + \alpha x_2)^T\|_E^2 \\ &\leqq (\|M_1 - M_2\|_{\sup} \|y(t) + \alpha x_2\| + \alpha \|M_1\|_{\sup} \|x_1 - x_2\|)^2 \\ &\quad + \sigma^2 \alpha^2 (\|y(t) + \alpha x_1\| + \|y(t) + \alpha x_2\|)^2 \|x_1 - x_2\|^2 \\ &\leqq 4\|M_1 - M_2\|_E^2 \|y(t) + \alpha x_2\|^2 \\ &\quad + [4\alpha^2 \|M_1\|_E^2 + \sigma^2 \alpha^2 (\|y(t) + \alpha x_1\| + \|y(t) + \alpha x_2\|)^2] \|x_1 - x_2\|^2 \end{aligned} \tag{2.5.6}$$

where we have used the inequality:

$$(a + b)^2 \leqq 4(a^2 + b^2). \tag{2.5.7}$$

Assume $|\alpha| \leqq 1$. Let $N = \frac{\sigma}{4\sqrt{2}}$. If $\|y(t)\| \leqq N$ $t \geqq 0$, and $\|(x, M)\| \leqq N$, then from (2.5.5):

$$\|F(t, (x, M))\|^2 \leqq \left(\frac{\sigma^2}{32} + \frac{\sigma^4}{8}\right) 4N^2 = \left(\frac{1}{8} + \frac{\sigma^2}{2}\right) \sigma^2 N^2 < N^2(1 - \|A\|_{\sup})^2 \tag{2.5.8}$$

and from (2.5.6):

$$\begin{aligned} &\|F(t, (x_1, M_1) - F(t, (x_2, M_2))\|^2 \\ &\leqq \frac{\sigma^2}{2} \|M_1 - M_2\|_E^2 + \left(\frac{\sigma^2}{8} + \frac{\sigma^4}{2}\right) \|x_1 - x_2\|^2 \\ &\leqq \tfrac{5}{8} \sigma^2 \|(x_1, M_1) - (x_2, M_2)\|^2 \end{aligned} \tag{2.5.9}$$

so the constant F, in Theorem 1.4 can be $\sqrt{\tfrac{5}{8}}\sigma < 1 - \|A\|_{\sup}$.

The conditions of Theorem 1.4 are therefore satisfied. So, if $|\alpha| \leqq 1$ and the inputs are not greater than $N = \frac{\sigma}{4\sqrt{2}}$ in norm then a bounded solution, or a T-periodic solution if $y(t)$ is T-periodic will exist.

It is possible to increase N and still use the theorem. However, it can be seen from (2.5.6) that when $x_1 = x_2 = 0$, we find an upper bounded for N, namely σ.

Figure 6 has some orbits plotted of (2.1.5) for $n = 1$, when $y(t)$ is periodic with period 4. Each $y(t)$ is not less than σ. From this we feel it is possible to prove the existence of periodic solutions for less restricted inputs.

We can also see that the periodic solution in Fig. 6 appears to be a limit cycle, that is, solutions approach it. There seems to be no theory pertaining to limit cycles for discrete systems at the present time.

2.6. Random Input

Figure 7 contains orbits of solutions for $n = 1$ in which the inputs are $y + e(t)$ where the $e(t)$'s are independent and identically distributed random variable having mean 0. Some solutions tend to stay trapped in an area. This area contains the asymptotically stable point of the discrete system with constant input y.

It appears that the model has some stable properties when the inputs fluctuate around some fixed vector. The non-linearity of the system, however, makes it difficult to analyze. We will investigate properties of the model under assumptions that linearize (2.1.5). First we state the following theorem:

Theorem 2.6. Let A be an $n \times n$ matrix. Let $Y = \{y(t)\}_{t=0}^{\infty}$ be a sequence of independent and identically distributed random vectors in R^n with distribution function F, mean $E(y)$, and covariance matrix $\operatorname{cov}(y)$. If all eigenvalues of A lie within the unit circle, then the solution $x(t)$ to:

$$x(t+1) = Ax(t) + y(t), \; x(0) = x_0 \tag{2.6.1}$$

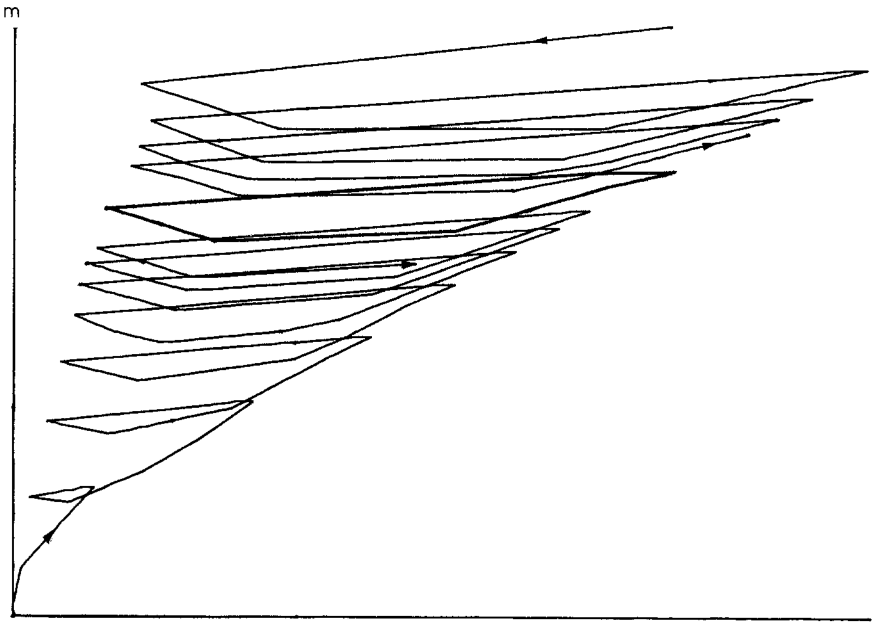

Fig. 6. Some orbits of solutions to (2.1.5) for $n=1$, with $\alpha=0.05$, $\varrho=0.9$ and $y(t)$ ranging periodically through (0.1, 0.5, 1.2, 1.7)

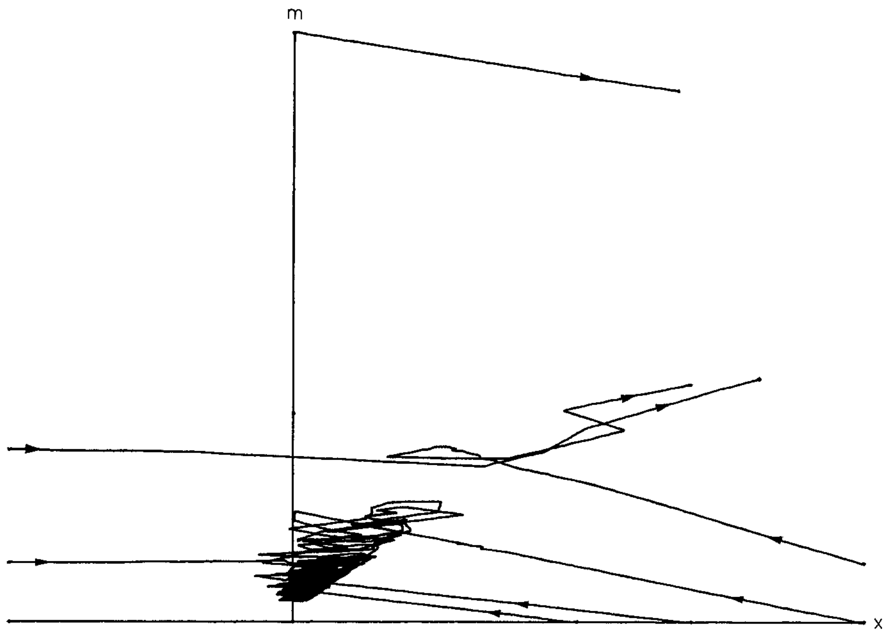

Fig. 7. Orbits of solutions to (2.1.5) for $n=1$, with $\alpha=0.2$, $\varrho=0.9$, and $y(t)=0.5+e(t)$ where the $e(t)$'s are independent and normally distributed random variables having mean 0 and variance 0.25

converges in distribution where the limiting distribution has mean:

$$(I-A)^{-1}E(y) \tag{2.6.2}$$

and covariance matrix:

$$\sum_{s=0}^{\infty} A^s \operatorname{cov}(y)(A^T)^s . \tag{2.6.3}$$

The theorem can be proven using characteristic functions.

At each time t the solution $x(t)$ to (2.6.1) is a random vector. The theorem only gives us conditions under which the distribution function of $x(t)$ will approach some fixed distribution. It is not true that $x(t)$ will converge to some random vector almost surely or in probability. But the limiting distribution can tell us in what regions realizations are likely to approach.

First we let $\varrho=1$. Then no learning will occur. If $(x(0), M(0))=(x_0, M_0)$ then:

$$x(t+1)=M_0(y(t)+\alpha x(t))=\alpha M_0 x(t)+M_0 y(t) . \tag{2.6.4}$$

If the eigenvalues of αM_0 are within the unit circle, then $x(t)$ converges in distribution with mean:

$$(I-\alpha M_0)^{-1}M_0E(y) \tag{2.6.5}$$

and covariance matrix:

$$\sum_{s=0}^{\infty} \alpha^{2s}M_0^{s+1}\operatorname{cov}(y)(M_0^T)^{s+1}. \tag{2.6.6}$$

We next assume $\alpha=0$, $\varrho<1$. We have

$$\begin{cases} x(t+1)=M(t)y(t) \\ M(t+1)=\varrho M(t)+\sigma y(t)y(t)^T. \end{cases} \tag{2.6.7}$$

Letting the i^{th} column of $M(t)$ to be $M_i(t)$ we write

$$\mathbf{x}(t)=\begin{bmatrix} M_1(t) \\ \vdots \\ M_n(t) \end{bmatrix} \tag{2.6.8}$$

$$\mathbf{y}(t)=\sigma\begin{bmatrix} y_1(t)y(t) \\ y_2(t)y(t) \\ \vdots \\ y_n(t)y(t) \end{bmatrix} \tag{2.6.9}$$

and get:

$$\mathbf{x}(t+1)=\varrho I_{n^2}\mathbf{x}(t)+\mathbf{y}(t). \tag{2.6.10}$$

If $\varrho<1$ then $\mathbf{x}(t)$ will converge in distribution with mean:

$$(I-\varrho I_{n^2})^{-1}E(\mathbf{y})=\frac{1}{\sigma}E(\mathbf{y}) \tag{2.6.11}$$

and covariance matrix:

$$\begin{aligned}\sum_{s=0}^{\infty}(\varrho I_{n^2})^s\operatorname{cov}(\mathbf{y})(\varrho I_{n^2})^s &= \frac{\sigma^2}{1-\varrho^2}\operatorname{cov}\begin{bmatrix} y_1y \\ \vdots \\ y_ny \end{bmatrix} \\ &= \frac{\sigma}{1+\varrho}\operatorname{cov}\begin{bmatrix} y_1y \\ \vdots \\ y_ny \end{bmatrix}. \end{aligned} \tag{2.6.12}$$

We conclude then that $M(t)$ converges in distribution with mean

$$E(yy^T). \tag{2.6.13}$$

We see from (2.6.12) that the limiting distribution can be centered as close to $E(yy^T)$ as desired by making σ small enough.

Finally, we linearize (2.1.5) for $\varrho<1$, $\alpha\neq0$. Let (x, M) be the critical point of (2.4.1) where $E(y)$ is substituted for y. Defining $x'(t)$, $M'(t)$ as in (2.4.18) the linear terms in the resulting equations are the same as those in (2.4.19). Therefore, if (x, M) satisfies the conditions for asymptotic stability in Theorem 2.4.3 ($\alpha>0$) or in Theorem 2.4.4 ($\alpha<0$), then the solutions to the linearized equations will converge in distribution with mean (x, M).

It seems reasonable to suggest that the full Eq. (2.1.5) behave almost the same way as the linearized ones near (x, M). But the non-linearity makes it very difficult to make any statement concerning the behavior of the random orbits. We feel, however, that the phenomenon observed in Fig. 7 can be understood in due time.

2.7. Remarks

We conclude from Section 2.4 that with constant input the network N, is capable of learning this input under the conditions implied by the definition of asymptotic stability and by Theorems 2.4.3 and 2.4.4. If $(x(0), M(0))$ is in the appropriate region in R^{n^2+n}, the matrix $M(t)$ will converge to the second term in (2.4.15), which is a multiple of the projection operator (defined in Section 2.3) associated with the input. This limiting matrix is uniquely determined by the input and the amplification factor α.

If the inputs are periodic with period T, then, under the conditions in Section 2.5, there will exist a T-periodic solution to (2.1.5). We predict this is the only T-periodic solution and will attract other solutions (i.e., is a limit cycle). This implies that the network, N, can learn (in the same sense as above) the T inputs since the T-periodic solution is uniquely determined by the inputs. However, we do not know how the solution is related to the inputs.

We also predict that inputs with small random fluctuations will result in a small region in R^{n^2+n} where most realizations of solutions starting near the region will eventually be trapped. This implies that N can learn approximately the idealized input, that is, the average value of the inputs.

We have assumed linear processing in discrete time intervals of neuronal spiking frequencies. Other recently proposed models differ from ours in some important assumptions. For example, Grossberg [9, 10] considers the processing in continuous time of membrane potentials. In his formulation the derivative with respect to time of the activity in the neurons at time t depends nonlinearly on the activity of time $t-\tau$. This results in a system of non-linear differential-difference equations, which are considerably more complex than our equations.

Processing in discrete time intervals was assumed mainly for convenience in computer simulations. We agree that in some respects a continuous time model is more realistic then a discrete time model. However, it is often computationally and conceptually easier to formulate the dynamics of processing in the form of an integral equation rather than a differential equation. Examples of recent related neuronal models using integral equations can be seen in Anderson [7] and Wilson and Cowan [21].

The inputs $y(t)$ can be considered to be neural representations of stimuli from a highly structured environment. We intend to use Pattern Analysis to model the environment and to view the network as a pattern processor.

One of our goals is to investigate properties of several interconnected networks. The understanding of one network is of course a prerequisite for the investigation of multi-layered models.

We must conclude that asymptotic analysis of the system with constant input will not, unfortunately, aid us in a more complete understanding of the learning properties of the model. It is clear that if the network receives one input long enough, the orbit of the system will, at best, approach the critical point associated with the input. The memory of all previous inputs will be wiped out. We should not deal, then, in long time spans. Instead, we should work in lengths in which the influence of the forgetting coefficient is not strong enough to remove earlier input.

The next step is to simulate learning experiences. Using the norm on R^n, a criteria for determining whether inputs are stored or not could be proposed. Combinations of inputs could be tested, under different choices of the parameters. New properties might be revealed using this approach which will lead to some interesting mathematics.

Acknowledgements. The author would like to thank Professors Ulf Grenander, James Anderson, E. F. Infante, Dr. Menasche Nass, and Mr. William Immerman for their invaluable suggestions and criticisms.

This work has been supported by the National Science Foundation under Grant GJ-31007X.

References

1. Deutsch, J. A.: The cholinergic synapse and the site of memory. Science **174**, 788—794 (1971)
2. Eccles, J. C.: Possible synaptic mechanisms subserving learning. In: Karczmar, A. G., Eccles, J. C. (Eds.): Brain and human behavior, pp. 39—61. Berlin-Heidelberg-New York: Springer 1972
3. Greenough, W. T.: Experiential modification of the developing brain. Amer. Sci. **63**, 37—46 (1975)
4. Horn, G., Rose, S. P. R., Bateson, P. P. G.: Experience and plasticity in the central nervous system. Science **181**, 506—514 (1973)
5. Anderson, J. A.: Two models for memory organization using interactive traces. Math. Biosci. **8**, 137—160 (1970)
6. Anderson, J. A.: A simple neural network generating an interactive memory. Math. Biosci. **14**, 197—220 (1972)
7. Anderson, J. A.: What is a distinctive feature? Techn. Rep. 74-1. Providence: Center for Neural Studies, Brown University 1974
8. Cooper, L. N.: A possible organization of animal memory and learning. In: Lundquist, B., Lundquist, S. (Eds.): Proceedings of the Nobel symposium on collective properties of physical systems. New York: Academic Press 1973
9. Grossberg, S.: Non-linear difference-differential equations in prediction and learning theory. Proc. Nat. Acad. Sci. U.S.A. **58**, 1329—1334 (1967)
10. Grossberg, S.: Embedding fields. A theory of learning with physiological implications. J. math. Psychol. **6**, 209—239 (1969)
11. Lashley, K. S.: In search of the engram. Symp. Soc. exp. Biol. **4**, 454—482 (1950)
12. Longuet-Higgins, H. C.: Holographic model of temporal recall. Nature **217**, 104 (1968)
13. Willshaw, D. J., Buneman, O. P., Longuet-Higgins, H. C.: Non-holographic associative memory. Nature (Lond.) **222**, 960—962 (1969)
14. McGaugh, J. M., Herz, J. J.: Memory consolidation. San Francisco: Albion Publ. Comp. 1972
15. Mountcastle, V. B.: The problem of sensing and the neural coding of sensory events. In: Quarton, G. C., Melnechuk, T., Schmitt, F. O. (Eds.): The neurosciences: A study program, pp. 393—408. New York: Rockefeller University Press 1967
16. Pfaff, D.: Parsimonius biological models of memory and reinforcement. Psychol. Rev. **76**, 70—81 (1969)
17. Grenander, U., Silverstein, J.: Spectral analysis of networks with random topologies. (submitted to SIAM J. appl. Math.)
18. Hahn, W.: Stability of motion. Berlin-Heidelberg-New York: Springer 1967
19. Hurt, J.: Some aspects of difference equations. Ph. D. Thesis. Providence: Brown University, Div. Appl. Math. 1967
20. Hurt, J.: Some stability theorems for ordinary difference equations, Tech. rep. no. 66-6. Providence: Center for Dynamical Systems, Div. Appl. Math., Brown University 1966
21. Wilson, H. R., Cowan, J. D.: A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. Kybernetik **13**, 55—80 (1973)

Dr. J. W. Silverstein
Division of Applied Mathematics
and Center for Neural Studies
Brown University
Providence, Rhode Island