A Study of the Asymptotic Behavior of Neural Networks

NIKITAS J. DIMOPOULOS, MEMBER, IEEE

Abstract —The stability properties of neural networks modeled as a set of nonlinear differential equations of the form $T\dot{X}+X=W\cdot f(X)+b$ where X is the neural membrane potential vector, W is the network connectivity matrix, and $f(X)$ is the nonlinearity (an essentially sigmoid function), are studied.

This paper establishes topologies of neural networks that exhibit asymptotic behavior. This behavior depends solely on the topology of the network. Moreover, the connectivity matrix W need not be symmetric.

Networks topologically similar to the cerebellum fall in this category and exhibit asymptotic behavior.

Simulated behavior of typical neural networks are also be presented.

I. Introduction

NEURAL NETWORKS have been studied for several decades now, with the recent flury of activity corresponding to the ability of constructing hardware analogs capable of mimicking aspects of the behavior of the central nervous system. Thus Mead [16] has constructed a 48×48 electronic analog of the retina that is capable of calculating time derivatives and hence model cells that look for motion. Psaltis *et al.* [19], [20], have implemented learning networks using holograms. Newcomb *et al.* [15], [14] have produced circuits that capture the temporal and spatial activity of multi-input neural elements, while Moopenn *et al.* [22], [17] have constructed electronic analogs of associative memory.

Several studies have been undertaken in the past years attempting to analyze the behavior of systems comprising a large number of interrelated and interacting elementary neurons. Thus Morishita *et al.* [18], [23] have studied networks of mutually inhibiting neurons. Hopfield [11]–[13] has studied the collective properties of systems consisting of neurons with binary or graded response. Properties such as absolute stability were proven for networks having symmetric interconnection matrices. Cohen and Grossberg [5] have also studied the absolute stability of global pattern formation in competitive neural networks. Again a prerequisite of absolute stability for such systems is the symmetry of the interconnection matrix. In [6], [7], the stability of neural networks topologically similar to the cerebellum was studied. In this work, we examine the stability properties of a generalized class of neural networks that contains networks that are topologically similar to the cerebellum. Such networks are composed of distinct classes of neurons and the manner that these classes interconnect provides for stability.

Manuscript received August 1, 1988; revised November 4, 1988. This work was supported by the Natural Science and Engineering Research Council of Canada under Grant A1337. This paper was recommended by Guest Editors R. W. Newcomb and N. El-Leithy.

The author is with the Department of Electrical and Computer Engineering, University of Victoria, Victoria, BC, Canada V8W 2Y2.

IEEE Log Number 88126715.

II. Description of the Network

We are interested in neural networks composed of k classes of neurons, each class having n_i; $i=1,2,\cdots,k$ neurons and adhering to the microscopic and macroscopic connectivity principles as discussed in [8]. The principles essentially state that a neural network is composed of a distinct and limited number of classes of similar neurons, and that neurons from a given class interconnect with neurons of other classes in a predetermined way that is invariant across a species (macroscopic connectivity principle). The strength of these interconnections vary from member to member, and they are subject to learning. Neurons connect to a small percentage of neighboring neurons, subject to the macroscopic connectivity principle (microscopic connectivity principle).

Such neural networks are described by systems of differential equations of the form:

$$T\dot{X}+X=W\cdot f(X)+b \tag{2.1}$$

where

$$\begin{aligned}X&=[X_1X_2\cdots X_k]^T\\&=[x_1x_2\cdots x_{n_1},x_{n_1+1},\cdots x_{n_1+n_2},\\&\qquad\cdots x_{N-n_{k-1}+1},\cdots,x_N]^T\end{aligned} \tag{2.2}$$

is the neural membrane potential vector,

$$N=\sum_{i=1}^{k}n_i \tag{2.3}$$

is the total number of neurons in the network

$$W=\begin{bmatrix}W_{11} & W_{12} & \cdots & W_{1k}\\ \vdots & & & \\ W_{k1} & W_{k2} & \cdots & W_{kk}\end{bmatrix} \tag{2.4}$$

is the network connectivity matrix.

$f(X)$ belongs to the class of the neuromime functions to be defined in Section III (c.f. Definition 3.2):

$$T = \operatorname{diag}[\tau_i], \qquad i = 1, 2, \cdots, N \tag{2.5}$$

is the membrane time constant, a positive matrix and finally $\boldsymbol{b}$ is a constant vector.

Of particular importance in formulating the system of differential equations (2.1), are the connectivity matrix W and the set of neuromime functions.

The connectivity matrix W, reflects the topology of the neural network in question. Each one of the submatrices W_{ij} may be either positive, negative or zero depending on whether neurons from class j excite, inhibit or have no direct effect whatsoever on neurons in class i.

As an example to clarify the point, take the case of networks topologically similar to the cerebellum [9], [1]–[3]. Such networks are comprised of four neural classes $\aleph_1$, $\aleph_2$, $\aleph_3$ and $\aleph_4$ corresponding to Purkiuje, basket, Golgi and granule cells. $\aleph_1$, $\aleph_2$, and $\aleph_3$ are classes of inhibitory neurons while $\aleph_4$ is a class of excitatory ones.

The connectivity matrix of such a network can be easily established from the organization of the cerebellum [9] and has the form

$$W = \begin{bmatrix} 0 & W_{12} & 0 & W_{14} \\ 0 & 0 & 0 & W_{24} \\ 0 & 0 & 0 & W_{34} \\ 0 & 0 & W_{43} & 0 \end{bmatrix} \tag{2.6}$$

with $W_{12}, W_{43} < 0$ and $W_{14}, W_{24}, W_{34} > 0$. This reflects the fact that, for example, the basket cells inhibit the Purkiuje cells ($W_{12} < 0$) and that there is no connection between basket and Golgi cells ($W_{23} = W_{32} = 0$).

The neuromime functions, on the other hand, describe the frequency behavior of the spike train on the axon as a function of the hillock potential. Generally speaking, spikes appear on the axon once the hillock potential exceeds a certain threshold [10], and their frequency increases with the potential. The neuromime functions, as defined in Section III incorporate these properties, namely they are positive non-decreasing functions with a threshold.

In the following, we shall examine the stability properties of neural networks described by (2.1). As it turns out, neural structures, such as the cerebellum, found in the central nervous system of several species, can be modelled effectively by such networks. One of the goals of our analysis is to establish fundamental properties of such structures that contribute to the asymptotic behavior. We were able to define a general class of such networks which we denote by $\wp$ and which exhibit asymptotic behavior. Neural networks modeling the cerebellum is but one element of this class.

III. Neural Networks Exhibiting Asymptotic Behavior

Definition 3.1: For every $(X, Y) \in R^N \times R^N$ with $x_i \leqslant y_i$ we say that vector X is less than or equal to vector Y and write $X \leqslant Y$.

Definition 3.2: The class $\aleph$ (the neuromime functions) is defined as follows:

$$\aleph = \{ f \mid f \colon R^N \to R_+^M,\ f \text{ continuous},\ f \text{ monotonically non-decreasing, satisfying a Lipschitz condition and } \exists \theta \in R^N \text{ such that } f(\theta) = 0 \}.$$

The postulation of Lipschitzian behavior for the neuromime functions, according to Definition 3.2, guarantees the uniqueness of solutions for the system of differential equations (2.1).

Definition 3.3: Given a vector in R^N, $X = [x_1, x_2, \cdots, x_N]^T$, we define $R_+^N \ni |X| = [|x_1|, |x_2|, \cdots, |x_n|]^T$, as the vector absolute value of the vector X.

Definition 3.4: Given the vector polynomial:

$$P(t) = \sum_{i=1}^{k} \boldsymbol{a}_i t^{i-1};\ \boldsymbol{a}_i \in R^N,\ t \in R_+$$

we define the vector absolute value of the polynomial $P(t)$ as

$$|P(t)| = \sum_{i=1}^{k} |\boldsymbol{a}_i| t^{i-1}.$$

Definition 3.5: A function $R^N \ni f(t, X)$ belongs to class ω if for every fixed $t \in R$ and all $X, Y \in R^N$ such that

$$x_j \leqslant y_j,\ x_i = y_i, \qquad j = 1, 2, \cdots, N;\ i \neq j$$

the following inequality is satisfied:

$$f_i(t, X) \leqslant f_i(t, Y) \quad \forall i = 1, 2, \cdots, N.$$

Definition 3.6: We will say that a system $\dot{X} = F(X, t)$ exhibits a poly-exponential behavior or equivalently belongs to class $\wp$ iff its solutions $X(t;\ X_0, t_0)$ are bounded in the following way:

$$K_* + e^{-A_* t} P_*(t) \leqslant X(t;\ X_0, t_0) \leqslant e^{-A^* t} P^*(t) + K^*$$

where A_*, A^* are positive diagonal matrices, P_*, P^* vector polynomials, K_*, K^* constant vectors, and $t \in R_+$.

Definition 3.7: We define the class $\aleph^*$ of the separable neuromime functions as

$$\aleph^* = \Big\{ f \mid f \in \aleph \text{ and } f(X) = [f_1(x_1), f_2(x_2) \cdots f_N(x_N)]^T \Big\}.$$

Lemma 3.1: For every function $g(t)$ such that

$$e^{-A_* t} P_*(t) + K_* \leqslant g(t) \leqslant e^{-A^* t} P^*(t) + K^* \tag{3.1.1}$$

then

$$e^{-A_* t} \underset{\sim}{P}(t) + \underset{\sim}{K} \leqslant f[g(t)] \leqslant e^{-A^* t} \tilde{P}(t) + \tilde{K} \tag{3.1.2}$$

where A_*, A^* are positive diagonal matrices $P_*(t)$, $P^*(t)$, $\underset{\sim}{P}(t)$, $\tilde{P}(t)$ vector polynomials K_*, K^*, $\underset{\sim}{K}$, $\tilde{K}$ constant vectors and $f \in \aleph$.

Proof: Certainly there exists an $M \in R^N$ such that

$$g(t) \leqslant M \quad \forall t \in R_+ \tag{3.1.3}$$

Then for every function $f \in \aleph$ we have

$$\left.\begin{matrix} \underline{H}X+\underline{K}, & \text{if } X \leqslant M \\ f(X), & \text{otherwise} \end{matrix}\right\} \leqslant f(X)$$

$$\leqslant \begin{cases} \overline{H}X+\overline{K}, & \text{if } X \leqslant M \\ f(X), & \text{otherwise} \end{cases} \quad (3.1.4)$$

with $\underline{H}$, $\overline{H} \geqslant 0$.

Then from (3.1.1), (3.1.3), and (3.1.4), we have

$$\underline{H}\left[e^{-A_*t}P_*(t)+K_*\right]+\underline{K} \leqslant f[g(t)]$$
$$\leqslant \overline{H}\left[e^{-A^*t}P^*(t)+K^*\right]+\overline{K} \quad (3.1.5)$$

or

$$e^{-A_*t}\underset{\sim}{P}(t)+\underset{\sim}{K} \leqslant f[g(t)] \leqslant e^{-A^*t}\tilde{P}(t)+\tilde{K} \quad (3.1.6)$$

where $\underset{\sim}{P}(t)=\underline{H}P_*(t)$, $\tilde{P}(t)=\overline{H}P^*(t)$

$$\underset{\sim}{K}=\underline{H}K_*+\underline{K} \text{ and } \tilde{K}=\overline{H}K^*+\overline{K}. \qquad \text{Q.E.D.}$$

Lemma 3.2: For every function $f \in \aleph$, every vector polynomial $P(t)$, vector K, and all positive diagonal matrices A and B there exist a vector polynomial $\hat{P}(t)$, a positive diagonal matrix $\hat{A}$, and a vector $\hat{K}$ such that the following is true:

$$0 \leqslant e^{-At}\int_0^t e^{As}f\left[e^{-Bs}P(s)+K\right]ds = I(t) \leqslant e^{-\hat{A}t}\hat{P}(t)+\hat{K}.$$

Proof: For a given $\theta \in R^N$:

(1) Definition *dl* 3.2.1: $R^N_{-\theta}=\{X|X \in R^N \text{ and } X<\theta\}$;

(2) Definition *dl* 3.2.2: $R^N_{+\theta}=\{X|X \in R^N \text{ and } X \geqslant \theta\}$;

(3) Definition *dl* 3.2.3: For a $K \in R^N_{+\theta}$ we symbolize by D_K a connected bounded subset of $R^N_{+\theta}$ such that $K \in D_K$.

Then $\forall D_\theta \subseteq R^N_{+\theta}$, $\forall f \in \aleph$, there exists an $H>0$ and a $\Lambda \in R^N$ such that

$$0 \leqslant f(X) \leqslant \begin{cases} 0 & \text{if } X \in R^N_{-\theta} \\ HX+\Lambda & \text{if } X \in D_\theta \\ f(X) & \text{otherwise.} \end{cases} \quad (3.2.1)$$

We have for $t \geqslant 0$

$$0 \leqslant I(t) \leqslant e^{-At}\int_0^t e^{As}f\left(e^{-Bs}|P(s)|+K\right)ds. \quad (3.2.2)$$

We recognize now two cases.

Case i): If $K<\theta$ then there exists a $\tilde{T}_0>0$ such that

$$e^{-Bt}|P(t)|+K \in R^N_{-\theta} \quad \forall t \geqslant \tilde{T}_0. \quad (3.2.3)$$

Therefore, the integral in (3.2.2) can be broken down as follows:

$$I(t) \leqslant e^{-At}\left[\int_0^{\tilde{T}_0} e^{As}f\left(e^{-Bs}|P(s)|+K\right)ds\right]$$
$$+e^{-At}\left[\int_{\tilde{T}_0}^t e^{As}f\left(e^{-Bs}|P(s)|+K\right)ds\right]. \quad (3.2.4)$$

But the first integral in (3.2.4) is bounded by a non-negative vector $\tilde{P}_0$ while the second, because of (3.2.3) and (3.2.1) is identically zero. Thus 3.2.4 gives

$$I(t) \leqslant e^{-At}\tilde{P}_0. \quad (3.2.5)$$

Case ii): If there exist some $i \in [1,\cdots,N]$ such that $k_i \geqslant \theta_i$, then we can write for (3.2.2)

$$0 \leqslant I(t) \leqslant e^{-At}\int_0^t e^{As}f\left(e^{-Bs}|P(s)|+K^*\right)ds \quad (3.2.6)$$

$$k_i^*=\begin{cases} k_i, & \text{if } k_i \geqslant \theta_i \\ \theta_i, & \text{if } k_i<\theta_i \end{cases}; \qquad K^*=\left[k_1^*,k_2^*,\cdots,k_N^*\right]^T. \quad (3.2.7)$$

Then for every D_{K^*} that includes θ there exists a $T_0 \geqslant 0$ so that

$$e^{-Bt}|P(t)|+K^* \in D_{K^*} \quad \forall t \geqslant T_0.$$

Accordingly, the integral in (3.2.6) can be broken down as follows:

$$I(t) \leqslant e^{-At}\left[\int_0^{T_0} e^{As}f\left(e^{-Bs}|P(s)|+K^*\right)ds\right.$$
$$\left.+\int_{T_0}^t e^{As}f\left(e^{-Bs}|P(s)|+K^*\right)ds\right]. \quad (3.2.8)$$

The first integral in (3.2.8) is bounded by a non-negative vector P_0, and taking into account (3.2.1) we have for (3.2.8)

$$I(t) \leqslant e^{-At}\left[P_0+\int_{T_0}^t e^{As}\left\{H\cdot\left(e^{-Bs}|P(s)|+K^*\right)+\Lambda\right\}ds\right] \quad (3.2.9)$$

or

$$I(t) \leqslant e^{-At}\hat{P}_0+H\cdot(A-B)^{-1}e^{-Bt}\tilde{P}(t)$$
$$+A^{-1}(H\cdot K^*+\Lambda) \quad (3.2.10)$$

where

$$\int e^{(A-B)s}|P(s)|\,ds=(A-B)^{-1}e^{(A-B)s}\tilde{P}(s).$$

$\tilde{P}(s)$ a vector polynomial in s and

$$\hat{P}_0=\left|P_0-H\cdot(A-B)^{-1}e^{(A-B)T_0}\tilde{P}(T_0)\right.$$
$$\left.-A^{-1}e^{AT_0}(H\cdot K^*+\Lambda)\right|. \quad (3.2.11)$$

Choosing now $\hat{B}=\min[A,B]$ with the meaning

$$\hat{B}_{ij}=\min\left[A_{ij},B_{ij}\right] \quad (3.2.12)$$

$$\hat{P}(t)=\hat{P}_0+|H\cdot(A-B)^{-1}\tilde{P}(t)| \text{ and}$$
$$\hat{K}=A^{-1}(H\cdot K^*+\Lambda) \quad (3.2.13)$$

we have

$$I(t) \leqslant e^{-\hat{B}t}\hat{P}(t)+\hat{K}. \quad (3.2.14)$$

Relations (3.2.5) and (3.2.14) prove the lemma. Q.E.D.

Remark R3.1: Observe that if $K=0$ and the threshold $\theta>0$ then Case i). holds and therefore, $I(t)$ goes asymptotically to 0 as $t \to \infty$.

Lemma 3.3: The system

$$\dot{X}+TX=-f(X)+g(t)+K \tag{3.3.1}$$

exhibits poly-exponential behavior (i.e., belongs to the class $\wp$) when $f(x)\in\aleph$,

$$K_*+e^{-A_*t}P_*(t)\leqslant g(t)\leqslant e^{-A^*t}P^*(t)+K^* \tag{3.3.2}$$

and T, A_*, A^* diagonal positive matrices $P_*(t)$, $P^*(t)$ vector polynomials and K, K^*, K_* constant vectors.

Proof: In this proof we make use of the comparison principle [21] and of Lemma 3.2. Since $f(x)\in\aleph$ then, by using (3.3.2) we have for (3.3.1):

$$\dot{X}+TX\leqslant g(t)+K\leqslant e^{-A^*t}P^*(t)+K+K^*. \tag{3.3.3}$$

Hence, by the comparison principle, the solution of (3.3.1) satisfies the following inequality:

$$X(t;X_0,t_0)\leqslant\rho(t;X_0,t_0) \tag{3.3.4}$$

where $\rho(t;X_0,t_0)$ is the solution of

$$\dot{\rho}+T\rho=e^{-A^*t}P^*(t)+\Lambda^* \tag{3.3.5}$$

passing through the point (X_0,t_0) and $\Lambda^*=K+K^*$. But then from (3.3.5) we have

$$\rho(t;X_0,t_0)=e^{-Tt}X_0+e^{-Tt}\int_0^t e^{Ts}\left[e^{-A^*s}P^*(s)+\Lambda^*\right]ds \tag{3.3.6}$$

or carrying on the integration in (3.3.6) we have

$$\rho(t;X_0,t_0)=e^{-Tt}\left(X_0-\tilde{P}(0)-T^{-1}\Lambda^*\right)+e^{-A^*t}\tilde{P}(t)+T^{-1}\cdot\Lambda^* \tag{3.3.7}$$

where

$$\int e^{(T-A)s}P^*(s)\,ds=e^{(T-A)s}\tilde{P}(s)$$

and $\tilde{P}(s)$ a vector polynomial or

$$\rho(t;X_0,t_0)\leqslant e^{-Tt}|X_0-\tilde{P}(0)-T^{-1}\Lambda^*|+e^{-A^*t}|\tilde{P}(t)|+T^{-1}\Lambda^*. \tag{3.3.8}$$

Setting $\hat{A}=\min[A^*,T]$, $\hat{K}=T^{-1}\cdot\Lambda^*$ and

$$\hat{P}(t)=|X_0-\tilde{P}(0)-T^{-1}\Lambda^*|+|\tilde{P}(t)|$$

we have from (3.3.8):

$$X(t;X_0,t_0)\leqslant\rho(t;X_0,t_0)\leqslant e^{-\hat{A}t}\hat{P}(t)+\hat{K}. \tag{3.3.9}$$

Using now (3.3.9) and the original system we have

$$\dot{X}+TX\geqslant-f\left[e^{-\hat{A}t}\hat{P}(t)+\hat{K}\right]+e^{-A_*t}P_*(t)+\Lambda_*$$
$$\Lambda_*=K_*+K. \tag{3.3.10}$$

Then by the comparison principle the solutions of (3.3.1) satisfy

$$X(t;X_0,t_0)\geqslant\nu(t,X_0,t_0) \tag{3.3.11}$$

where $\nu(t;X_0,t_0)$ is the solution of

$$\dot{\nu}+T\nu=-f\left[e^{-\hat{A}t}\hat{P}(t)+\hat{K}\right]+e^{-A_*t}P_*(t)+\Lambda_* \tag{3.3.12}$$

passing through the point (X_0,t_0). Integrating (3.3.12), we have

$$\nu=e^{-Tt}X_0-e^{-Tt}\int_0^t e^{Ts}f\left[e^{-\hat{A}s}\hat{P}(s)+\hat{K}\right]ds+e^{-Tt}\int_0^t e^{Ts}\left[e^{-A_*s}P_*(s)+\Lambda_*\right]ds.$$

Designating by

$$I_1(t)=-e^{-Tt}\int_0^t e^{Ts}f\left[e^{-\hat{A}s}\hat{P}(s)+\hat{K}\right]ds \tag{3.3.13}$$

and

$$I_2(t)=e^{-Tt}\int_0^t e^{Ts}\left[e^{-A_*s}P_*(s)+\Lambda_*\right]ds \tag{3.3.14}$$

we have from (3.3.12)

$$\nu=e^{-Tt}X_0+I_1(t)+I_2(t). \tag{3.3.15}$$

But by Lemma 3.2:

$$I_1(t)\geqslant-e^{-{}_*At}{}_*P(t)+{}_*K \tag{3.3.16}$$

where ${}_*A$, ${}_*P(t)$, and ${}_*K$ are calculated for T, $\hat{A}$, $\hat{P}(t)$, and $\hat{K}$ following the procedure outlined in Lemma 3.2. Carrying out the integration in (3.3.14) we have

$$I_2(t)=e^{-A_*t}\tilde{P}_*(t)-e^{-Tt}\left[\tilde{P}_*(0)+T^{-1}\Lambda_*\right]+T^{-1}\Lambda_* \tag{3.3.17}$$

where

$$\int e^{(T-A_*)s}P_*(s)\,ds=e^{(T-A_*)s}\tilde{P}_*(s)$$

and $\tilde{P}_*(s)$ a vector polynomial in s or

$$I_2(t)\geqslant-e^{-A_*t}|\tilde{P}_*(t)|-e^{-Tt}|\tilde{P}_*(0)+T^{-1}\Lambda_*|+T^{-1}\Lambda_*. \tag{3.3.18}$$

By denoting

$$\bar{A}=\max[A_*,T];\ \bar{P}(t)=|\tilde{P}_*(t)|+|\tilde{P}_*(0)+T^{-1}\Lambda_*|;$$
$$\bar{K}=T^{-1}\Lambda_* \tag{3.3.19}$$

we have $I_2(t)\geqslant-e^{-\bar{A}t}\bar{P}(t)+\bar{K}$.

Combining (3.3.11), (3.3.16), and (3.3.19) we have

$$X(t;X_0,t_0)\geqslant-e^{-{}_*At}{}_*P(t)+{}_*K-e^{-\bar{A}t}\bar{P}t)+\bar{K}. \tag{3.3.20}$$

Choosing

$$A=\max\left[{}_*A,\bar{A}\right],\ P(t)={}_*P(t)+\bar{P}(t);\ K={}_*K+\bar{K} \tag{3.3.21}$$

we have

$$X(t;X_0,t_0)\geqslant-e^{-At}P(t)+K. \tag{3.3.22}$$

Relations (3.3.9) and (3.3.22) prove the lemma. Q.E.D.

Remark R3.2: Observe that if $K^*=K_*=K=0$ and the threshold θ in $f(x)$ is positive then K and $\hat{K}$ become zero and the solutions of the system (3.3.1) go to zero as $t\to\infty$.

Lemma 3.4: If the system

$$\dot{X}+TX=W\cdot f(X)+K \tag{3.4.1}$$

belongs to the class $\wp$ so does the system

$$\dot{X}+TX=W\cdot f(X)-g(X)+K \qquad (3.4.2)$$

with $f(X)$, $g(X)\in\aleph$; W the connectivity matrix, T a positive diagonal matrix, and $W\cdot f(X)\in\omega$.

Proof: For the system (3.4.2) we have

$$\dot{X}+TX\leqslant W\cdot f(X)+K \qquad (3.4.3)$$

Since $W\cdot f(X)\in\omega$ then $-TX+W\cdot f(X)+K\in\omega$. Therefore, by the comparison principle, the solutions of (3.4.2) will satisfy the condition

$$X(t;X_0,t_0)\leqslant\rho(t;X_0,t_0) \qquad (3.4.4)$$

where $\rho(t;X_0,t_0)$ is the solution of

$$\dot{\rho}+T\rho=W\cdot f(\rho)+K. \qquad (3.4.5)$$

But from the statement of the lemma, system (3.4.5) belongs to class $\wp$, hence

$$X(t;X_0,t_0)\leqslant\rho(t;X_0,t_0)\leqslant e^{-A_*t}P_*(t)+K_*. \qquad (3.4.6)$$

On the other hand, we have for the system (3.4.2)

$$\dot{X}+TX\geqslant W_-f(x)-g(X)+K=-\boldsymbol{h}(X)+K \qquad (3.4.7)$$

where W_- is the connectivity matrix W with all positive entries nullified. It is obvious that $\boldsymbol{h}(X)\in\aleph$. Using (3.4.6) and the fact that since $\boldsymbol{h}(X)\in\aleph$, $\boldsymbol{h}(X)$ is monotonically non-decreasing we have

$$\dot{X}+TX\geqslant-\boldsymbol{h}(X)+K\geqslant-\boldsymbol{h}\left[e^{-A_*t}P_*(t)+K_*\right]+K \qquad (3.4.8)$$

Observe now that system (3.4.8) is decoupled and Lipschitzian. Therefore, by the comparison principle the solutions of (3.4.8) satisfy the next inequality.

$$X(t;X_0,t_0)\geqslant\nu(t;X_0,t_0) \qquad (3.4.9)$$

where $\nu(t;X_0,t_0)$ is the solution of

$$\dot{\nu}+T\nu=-\boldsymbol{h}\left[e^{-A_*t}P_*(t)+K_*\right]+K. \qquad (3.4.10)$$

From (3.4.10) we have

$$\nu=e^{-Tt}\left(X_0-T^{-1}K\right)+T^{-1}K$$
$$-e^{-Tt}\int_0^t e^{Ts}\boldsymbol{h}\left[e^{-A_*t}P_*(t)+K_*\right]ds \qquad (3.4.11)$$

But by Lemma 3.2:

$$-e^{-Tt}\int_0^t e^{Ts}\boldsymbol{h}\left[e^{-A_*t}P_*(t)+K_*\right]ds\geqslant-e^{-{}_*At}{}_*P(t)+{}_*K \qquad (3.4.12)$$

where ${}_*A$, ${}_*P(t)$ and ${}_*K$ are calculated from T, A_*, $P_*(t)$, and K following the procedure of Lemma 3.2.

Therefore, combining (3.4.11) and (3.4.12) we have

$$\nu\geqslant e^{-Tt}\left(X_0-T^{-1}K\right)+T^{-1}K-e^{-{}_*At}$$

or

$$X(t;X_0,t_0)\geqslant\nu\geqslant-e^{-A^*t}P^*(t)+K^* \qquad (3.4.13)$$

with

$$A^*=\max\left[{}_*A,T\right],\qquad K^*=T^{-1}K+{}_*K$$

and

$$P^*(t)=\left(T^{-1}K-X_0\right)+{}_*P(t).$$

Relations (3.4.6) and (3.4.13) prove the lemma. Q.E.D.

Lemma 3.5: The system

$$\dot{X}+TX=W_p\cdot f(X)+K \qquad (3.5.1)$$

belongs to class $\wp$, when $f(X)\in\aleph^*$ and W_p is a connectivity matrix with all the elements above the diagonal non-negative and the rest zero.

Proof: System (3.5.1) can be broken down in the following form:

$$\begin{aligned}
\dot{x}_1+Tx_1&=W_{12}f_2(x_2)+W_{13}f_3(x_3)+\cdots+W_{1n}f_n(x_n)+K_1 && (3.5.2;1)\\
\dot{x}_2+Tx_2&=\phantom{W_{12}f_2(x_2)}+W_{23}f_3(x_3)+\cdots+W_{2n}f_n(x_n)+K_2 && (3.5.3;2)\\
&\vdots\\
\dot{x}_{n-1}+Tx_{n-1}&=W_{n-1,n}f_n(x_n)+K_{n-1} && (3.5.2;n\text{-}1)\\
\dot{x}_n+Tx_n&=K_n. && (3.5.2;n)
\end{aligned}$$

Starting with (3.5.2; n) we have

$$x_n=\left(x_{n0}-T^{-1}K_n\right)e^{-Tt}+T^{-1}K_n. \qquad (3.5.3;n)$$

From (3.5.3; n) and (3.5.2; n-1) we have

$$\dot{x}_{n-1}+Tx_{n-1}=W_{n-1,n}f\left[\left(x_{n_0}-T^{-1}K_n\right)e^{-Tt}+T^{-1}K_n\right]$$
$$+K_{n-1}. \qquad (3.5.3;n\text{-}1)$$

But by Lemma 3.1, $f[(x_{n_0}-T^{-1}K_n)e^{-Tt}+T^{-1}K_n]$ is bounded by functions of the form $e^{-Tt}P(t)+K$ and hence by Lemma 3.3 system (3.5.3; n-1) belongs by itself to class $\wp$.

Making now repeated use of Lemmas 3.1 and 3.2 and proceeding as above we can conclude that all systems (3.5.2; n-2)–(3.5.2; 1) belong to class $\wp$ or equivalently the original system (3.5.1) belongs to class $\wp$.

Theorem 3.1: The system

$$\dot{X}+TX=Wf(X)+K \qquad (\mathrm{T3.1.1})$$

with $f(x) \in \aleph^*$ and where the connectivity matrix W has all its positive entries located above the main diagonal belongs to class $\wp$.

Proof: The system (T3.1.1) can be written as

$$\dot{X} + TX = W_+ f(X) + W_- f(X) + K \quad \text{(T3.1.2)}$$

where $W_+ + W_- = W$ and W_+ contain all the positive elements of W while W_- all the negative ones. Zeros are filling up the other positions. Obviously, $W_+ f(x) \in \omega$ since $W_+ \geqslant 0$. Then by Lemma 3.5 the system

$$\dot{X} + TX = W_+ f(X) + K \quad \text{(T3.1.3)}$$

belongs to class $\wp$ and by Lemma 3.4 so does the system

$$\dot{X} + TX = W_+ f(X) + W_- f(X) + K. \quad \text{(T3.1.4)}$$

Q.E.D.

Theorem 3.2: Neural networks having the topology of the cerebellum belong to class $\wp$.

Proof: Neural networks with the cerebellum topology can be described as

$$\dot{X} + TX = Wf(X) + K \quad \text{(T3.2.1)}$$

where W has the form given in (2.6).

As we can see, the positive entries in this matrix are located above the main diagonal. Hence by Theorem 1 system (T3.2.1) belongs to class $\wp$.

IV. Discussion

The system studied in the previous paragraph represents neural nets where the nonlinearities of the individual neurons in the net are described by the neuromime functions of the class $\aleph$. Functions that belong to the class $\aleph$ are positive definite functions that exhibit a cutoff. The class $\aleph$ is a general class that accommodates many of the recently developed neuron models with a prime example the ones studied by Morishita [18], [23] and Heiden [10].

In that model the nonlinearity has the following form:

$$f_m(x) = x \cdot \phi[x]$$

where

$$\phi[X] = \begin{cases} 0, & \text{if } X \leqslant 0 \\ 1, & \text{if } X > 0 \end{cases}.$$

The class of neuromime functions $\aleph$, also contains the general sigmoid functions used by Hopfield *et al.* [11], [12], [13], and resemble the functions studied by Grossberg *et al.* [5]. It is worthwhile to note that in the calculus of Section III we only made use of the properties of the class $\aleph$ and, therefore, the results qualify for any specific function $f \in \aleph$ and depend only on the topology of the network in question.

The results obtained in this work constitute a generalization of the ones presented in [7] in the sense that they qualify a much broader category of neural networks, those with connectivity matrices having the form stated in Theorem 3.1, as exhibiting asymptotic behavior. It is important to point out that the class of the neuromime functions has also been enlarged to include functions with arbitrary

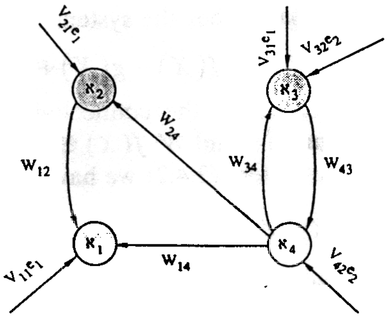

Fig. 1. Neural network topologically similar to the cerebellum. $\aleph_1$: Class of Purkinje cells (inhibitory); $\aleph_2$: Class of Basket cells (inhibitory); $\aleph_3$: Class of Golgi cells (inhibitory); $\aleph_4$: Class of granule cells (excitatory); e_1: Input from climbing fibers; e_2: Input from mossy fibers.

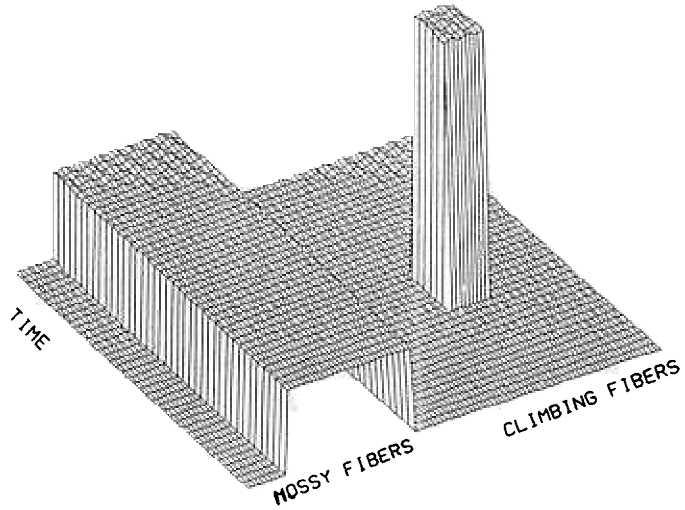

Fig. 2. The input to the simulated neural network.

thresholds. This was done in order to accommodate the modeling of neural networks containing classes of self-excitatory neurons. Such neurons may, for example, be found in the Hippocampus [4].

If on the other hand, the neuromime functions include only those with positive thresholds, then the constants K and Λ found in Lemmas 3.1–3.5 and Theorems 3.1 and 3.2, can be set to zero, resulting in the asymptotic behavior of such neural networks. An analysis of neural networks resulting from neuromime functions with positive thresholds, can be found in [7]. Certainly, as stated in Theorem 3.2, networks with the topology of the cerebellum fall to the same category and, therefore, exhibit asymptotic behavior.

We have also constructed a simulator, that provides us with a useful tool of studying the effect of the structure on the behavior of neural networks. In Figs. 2–6, we present the results obtained with the above mentioned simulator, for neural networks that are topologically similar to the cerebellum. This type of networks consist of four classes of neurons, with interconnections depicted in Fig. 1. The results presented here were obtained from a neural network comprising 200 neurons with randomly selected interconnection weights and thresholds. Thus Figs. 3–6 show

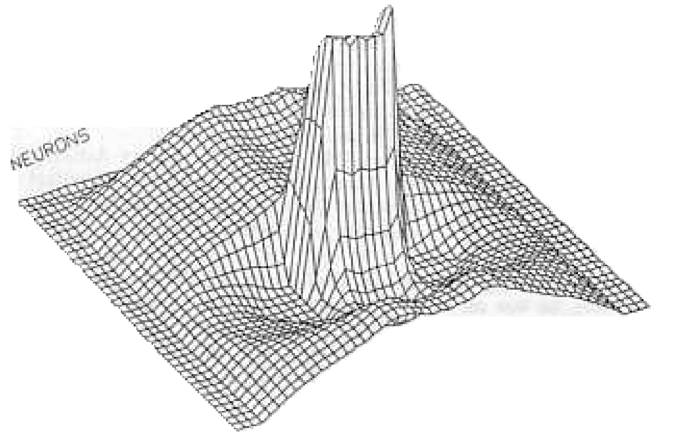

Fig. 3. Firing frequencies of the Purkinje cells (class $\aleph_1$).

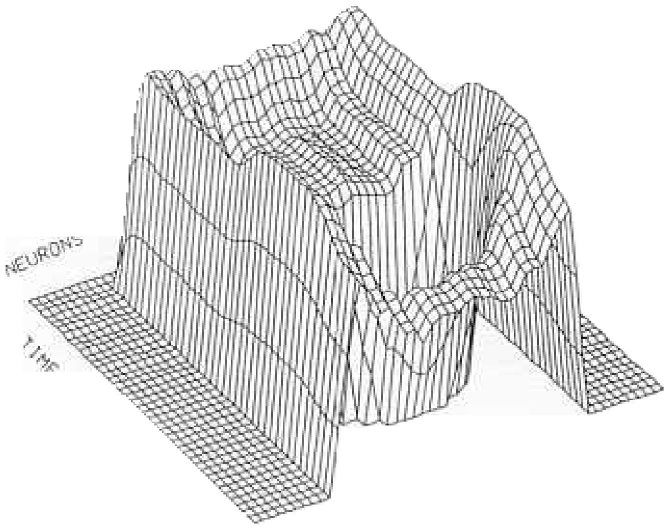

Fig. 4. Firing frequencies of the granule cells (class $\aleph_4$).

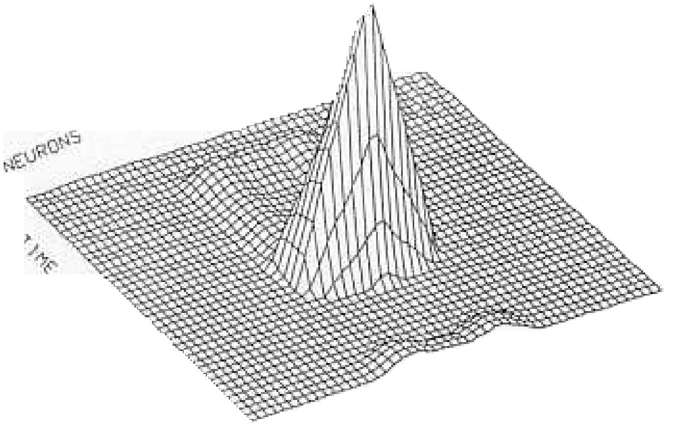

Fig. 5. Firing frequencies of the basket cells (class $\aleph_2$).

the firing frequencies of the various neural classes (Purkinje, granule, basket, and Golgi cells, respectively) as a result of external stimulation through the mossy fiber and climbing fiber systems. The external stimuli are shown in Fig. 2. It is apparent from the presented results the asymptotic behavior of such networks, as well as the existence of a complex spike as a response to a voley in the

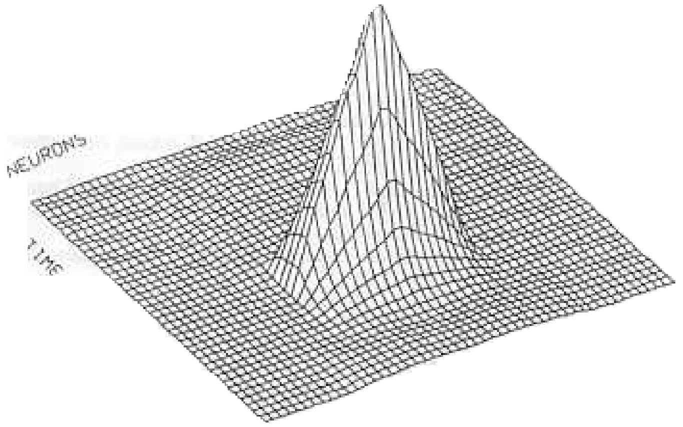

Fig. 6. Firing frequencies of the Golgi cells (class $\aleph_3$).

climbing fibers by the Purkinje cells. The complex spike comprises a sharp increase of activity in the Purkinje cells followed by a period of silence, and corresponds exactly to the physiological observations cited by Armstrong [3]. The silencing of the granule cells due to the enmaron climbing fibre volley is easily demonstrated in Fig. 4. This silencing of the granule cells was postulated by Armstrong [3] to be the cause of the Purkinje cell silencing after a complex spike.

References

[1] J. C. Eccles, M. Ito, and J. Szentaghothai, *The Cerebellum as a Neuronal Machine*. New York: Springer Verlag, 1967.

[2] D. Marr, "Atheory of cerebellar cortex," *J. Physiol.*, vol. 203, pp. 437–470, 1969.

[3] D. M. Armstrong, "The mamalian cerebellum and its contribution to movement control," in *Int. Rev. Physiology*. R. Porter, Ed., Univ. Park Press, Baltimore, pp. 239–293, 1978.

[4] P. Andersen "Organization of Hippocampal neurons and their interconnections," in *The Hippocampus*, R. Isaacson and K. Primram, Eds, vol. 1, Plenum, pp. 155–175, 1975.

[5] M. A. Cohen and S. Grossberg, "Absolute stability of global pattern formation and parallel memory storage by competitive neural networks," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-13, pp. 815–826, Sept./Oct. 1983.

[6] N. Dimopoulos and R. W. Newcomb, "Modeling networks of Morishita neurons with application to the cerebellum," in *Proc. Int. Conf. on Cybernetics and Society*, Denver, CO, pp. 597–602, Oct. 1979.

[7] ——, "Stability properties of a class of large scale neural networks," in *Proc. 1980 IEEE Int. Symp. on Circuits and Systems*, Houston, TX, pp. 528–530, Apr. 1980.

[8] N. Dimopoulos, "Organization and stability of a neural network class and the structure of a multiprocessor system," Ph.D. dissertation, Univ. of Maryland, 1980.

[9] J. C. Eccles. *The Understanding of the Brain*. New York: McGraw-Hill, 1973.

[10] U. an Der Heiden, "Structures of Excitation and Inhibition," in *Theoretical Approaches to Complex Systems*, S. Levine Ed., *Lecture Notes in Biometh.*, Springer-Verlag, vol. 21, pp. 75–88, 1977.

[11] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," in *Proc. Natl. Acad. Sci. USA*, vol. 79, pp. 2554–2558, Apr. 1982.

[12] J. J. Hopfield and D. W. Tank, "Computing with neural circuits: A model," *Science*, vol. 233, pp. 625–633, Aug. 1986.

[13] J. J. Hopfield, "Neurons with graded response have collective computational properties like those of two-state neurons," in *Proc. Natl. Acad. Sci. USA*, vol. 81, pp. 3088–3092, May 1984.

[14] C. K. Kohli, R. C. Ajmera, G. Kiruthi, and R. W. Newcomb, "Hysteretic system for neural-type circuits," *Proc. IEEE*, vol. 69, pp. 285–287, Feb. 1981.

[15] N. El-Leithy, R. W. Newcomb and M. Zaghloul, "A basic MOS neural-type junction. A perspective on neural type microsystems," in *Proc. IEEE First Int. Conference on Neural Networks*, vol. III, pp. 469–477, June 1987.

[16] Carver Mead, "Silicon models of neural computation," *Proc. IEEE First Int. Conf. on Neural Networks*, vol. I, pp. 91–106, June 1987.

[17] A. Moopenn, A. P. Thakoor, T. Duong, and S. K. Khanna, "A neurocomputer based on an analog-digital hybrid architecture," *Proceedings IEEE First Int. Conf. on Neural Networks*, vol. III, pp. 479–486, June 1987.

[18] I. Morishita and A. Yajima, "Analysis and simulation of networks of mutually inhibiting neurons," *Kyberuetik*, pp. 154–165, 1972.

[19] Y. S. Abu-Mostafa and D. Psaltis, "Optical neural computers" *Scientific Amer.* vol. 256, no. 3, p. 88, 1987.

[20] D. Psaltis, K. Wagner, and D. Brady, "Learning in optical neural computers," in *Proc. IEEE First Int. Conf. on Neural Networks*, vol. III, pp. 549–555, Jun. 1987.

[21] D. D. Siljak, *Large Scale Dynamic Systems*. Amsterdam, The Netherlands: North Holland, 1978.

[22] A. P. Thakoor, "Content-addressable, High density memories based on neural network models," Technical Progress Rep. JPL D-4166, California Inst. Technol., Pasadena CA, Mar. 1987.

[23] T. Tokura and I. Morishita, "Analysis and simulation of double-layer neural networks with mutually inhibiting interconnections," *Biol. Cybern.*, vol. 25, pp. 83–92, 1977.

Nikitas J. Dimopoulos (S'78–M'81) received the B.Sc. degree in physics from the University of Athens, Athens, Greece, in 1975, and the M.Sc. and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, MD. in 1976 and 1980, respectively.

From 1980 to 1988, he was an Assistant and then Associate Professor at the Electrical and Computer Engineering Department, Concordia University in Montreal. During the Summer of 1986 he was on leave with the Centre de Recherche Informatique de Montreal. From November 1986 to December 1987, he was on leave at the Jet Propulsion Laboratory, Pasadena, CA. In 1988, he joined the Electrical and Computer Engineering Department, University of Victoria, Victoria, BC where he is currently an Associate Professor. His research include multiprocessors, local area networks and neural networks.