Averaged Neural Networks

T.A. Burton
Southern Illinois University at Carbondale

Abstract: *Cohen and Grossberg (1983) studied an almost gradient system of ordinary differential equations with application to neural networks and used a Liapunov function, together with an invariance principle, to show that some equilibrium points attract solutions. Independently, Hopfield (1984) modeled a neural network by means of a system of ordinary differential equations which turn out to be a special case of the Cohen-Grossberg system, as pointed out by Cohen (1990). In the Hopfield model it is clear that the functions involved in the equations are averages and that current will flow through a synapse only if a certain threshold is reached; however, none of the models take into account an averaging technique.*

Investigators have been interested in sustained oscillations in neural networks and have produced them in computer simulations when there is a pointwise delay. The linearized systems with a delay have also exhibited sustained oscillations.

But our conjecture is that oscillations are not caused by a delay. This paper is intended to put substance to that conjecture by examining models with both pointwise and distributed delays. None of the models have solutions with sustained oscillations.

Keywords. Cohen-Grossberg system, Hopfield system, Liapunov functions, stability.

Acknowledgment. This work was supported in part by the Neural Engineering Research Center, Southern Illinois University, Carbondale, Illinois. Correspondence should be addressed to: T.A. Burton, Department of Mathematics, Southern Illinois University, Carbondale, Illinois 62901-4408.

1. INTRODUCTION

In this paper we study stability properties of several systems of differential equations which are intended to model certain neural networks. The equations have also been of considerable mathematical interest in the past.

The study begins with the system of Cohen and Grossberg (1983) and Hopfield (1984)

$$\dot{u}_i = -(\partial E(u)/\partial u_i)/g_i'(u_i), \quad i = 1, \ldots, n, \tag{1}$$

(where E is a given smooth scalar function and g_i is sigmoidal) which has been studied for many years. A careful stability analysis of a generalization of it was given by Cohen and Grossberg (1983) using a Liapunov function ($E(u)$) and an invariance principle. A well-motivated derivation of it for neural networks was given by Hopfield (1982 and 1984) who did not write it in such a compact form, but it is of this type as was pointed out by Cohen (1990).

In Hopfield's derivation two items are prominent:

First, each term in (1) is an approximation of a "short term average."

Next, u_i represents the charge on the ith neuron. Many other neurons connect to the ith one through a synapse. If the total charge at the synapse at a given time is below a threshold level, then the synapse does not fire. In fact, the charges collect at the synapse waiting for the threshold to be reached, but dissipate quickly (perhaps in 3 or 4 milliseconds (see Grundfest (1967)) if the sum of the charges does not reach the threshold level.

We first present a model which will follow (1) but which will take into account both the fact that the charge at the synapse is collected and averaged over a time period, and that it dissipates. The dissipation is modeled by selecting functions $a_i(t)$ satisfying

$$a_i(0) > 0, \quad a_i'(t) \le 0, \quad a_i(T) = 0, \quad a_i''(t) > 0 \tag{2}$$

where T is the amount of time it takes for a charge at the synapse to dissipate if the total charge does not reach the threshold level. Our model then is

$$\dot{u}_i = -[g_i'(u_i)]^{-1/2} \int_{t-T}^{t} a_i(t-s)[\partial E(u(s))/\partial u_i][g_i'(u_i(s))]^{-1/2} ds. \tag{3}$$

(The coefficient of the integral is placed there instead of inside as a technical necessity.) We will discuss (3) in more detail later, but point out here that the stability analysis turns out to be similar to that for (1); moreover, every solution of (3) approaches the equilibrium set of (1). Owing to the properties of the equilibrium points in neural networks, it is very important that those points be preserved in any model that purports to generalize (1). Discussions of equilibrium points are found in Cohen and Grossberg (1983) and in Hopfield (1984). System (3) is a generalization of a scalar model of Levin and Nohel (1964) and Hale (1977) (see p. 120) who also assumed (2) and used a sophisticated invariance principle in the stability analysis.

Investigators have conjectured that observed sustained oscillations in neural networks are caused by a delay. Marcus and Westervelt (1989a,b) discuss this extensively in terms of a pointwise delay which they believe (based on computer simulations and linearization arguments) point to sustained oscillations.

But we conjecture otherwise. In the actual systems, any delay is very short. System (1) is almost a gradient system, and gradient systems with asymptotically stable equilibrium points are the most strongly stable systems possible, as geometrical arguments will show. This means that small delays should not disrupt the asymptotic stability. But the structure of (1) is so strongly stable that when a delay is introduced in the nonlinear part of (1) and when the linear part dominates the nonlinear part, then there are no sustained oscillations for any size of the delay. This is the content of our second result.

2. THE EQUATION WITH MEMORY

Consider the system (1) in which E is a scalar function of u having the property that:

$$\left.\begin{array}{r}\text{For each } u^0 \in R^n, \text{ there is an } R > 0 \text{ such that} \\ |u| \geq R \text{ implies that } E(u) > E(u^0).\end{array}\right\} \tag{4}$$

When (4) holds, if $u(t)$ is a solution of (1) and if $\dot{E}(u(t)) \leq 0$, then $u(t)$ is bounded. We follow Hopfield and ask that

$$g_i'(r) > 0, \quad g_i(0) = 0, \tag{5}$$

while

$$\partial E/\partial u_i \text{ and } g_i' \text{ are continuous.} \tag{6}$$

Under these conditions, for each initial point $u^0 \in R$ there is at least one solution $u(t, u^0)$ of (1) on an interval $[0, \alpha)$ and, if the solution remains bounded, then $\alpha = \infty$. This is the classical Cauchy-Peano existence theorem (cf. Burton (1985), pp. 183–5). We first state the well-known Cohen and Grossberg (1983) and Hopfield (1984) theorem.

Theorem 1: Let (4), (5), and (6) hold for (1). Then every solution of (1) is bounded and approaches the set of equilibrium points of (1).

We turn now to (3) with (2) holding and offer a few words of motivation. Hopfield's system is

$$C_i \dot{u}_i = \sum_{j=1}^{n} T_{ij} V_j - u_i/R_i + I_i \tag{7}$$

where $V_j = g_j(u_j)$, g_j is sigmoidal, $g_j(0) = 0$, $g_j'(u_j) > 0$, $g_j(u_j) \to \pm 1$ as $u_j \to \pm\infty$. Now u_i is taken to be the mean potential of the ith neuron, while $V_j = g_j(u_j)$ is the input-output relation where the jth neuron connects to the ith neuron at a synapse, and T_{ij} is the efficacy of the synapse. Also, C_i is the capacitance of the cell membrane and I_i represents any other fixed input to the ith neuron.

When

$$E = -(1/2) \sum_i \sum_j T_{ij} V_i V_j + \sum_i (1/R_i) \int_0^{V_i} g_i^{-1}(s)\,ds - \sum_i I_i V_i \tag{8}$$

and when $T_{ij} = T_{ji}$, then it is readily seen, and noted by Cohen-Grossberg (1983) (p. 818) (see also Cohen (1990)), that (1) holds.

As mentioned in the introduction, if the charge at the synapse does not reach a certain threshold level, then it does not fire and it quickly dissipates. Thus, we want to explicitly average the right-hand side of (7), account for the dissipation of the charge at the synapse and in the neuron, and maintain the location of the equilibrium points of (1).

Let T be the time required for a charge to dissipate at the synapse. The average of a function q on an interval $[t-T,t]$ is $(1/T)\int_{t-T}^{t} q(s)ds$. But to let the charge dissipate in T time units according to a decay curve a(t), we form $\int_{t-T}^{t} a(t-u)q(u)du$ and note that when $u=t$ the integrand is $a(0)q(t)$, while at $u=t-T$ the integrand is $a(T)f(t-T)$.

A word about existence of solutions of (3) is in order and the reader is referred to Burton (1985) (see pp. 186–189) or to Hale (1977) (see p. 41) for details. To determine a solution of (3) we require a continuous initial function $\phi : [-T,0] \to R^n$ with $u(t)=\phi(t)$ on $[-T,0]$. By the continuity of the equation and the fact that bounded sets are mapped into bounded sets, one can prove that there is a solution $u(t)=u(t,\phi)$ on some interval $[0,\alpha)$. If the solution remains bounded then $\alpha=\infty$. We will define a Liapunov functional V along a solution of (3) with $\dot{V} \le 0$ which, together with our new assumption

$$(4^*) \qquad E(u) \to \infty \text{ as } |u| \to \infty,$$

will show that all solutions of (3) are bounded and, hence, continuable on $[0,\infty)$. Here, it will be convenient to explicitly use the fact that the right-hand-side of (3) is bounded for u bounded.

Theorem 2: Let (2), (4*), (5), and (6) hold for (3) with $T>0$. Then every solution of (3) is bounded and approaches the equilibrium set of (1).

Proof: Let $u(t)$ be any solution of (3) and for $u=u(t)$ define the functional

$$(9) \qquad V = 2E(u) - \sum_{i=1}^{n} \int_{t-T}^{t} a_i'(t-s)\left(\int_s^t \{(\partial E(u(v))/\partial u_i)[g_i'(u_i(v))]^{-1/2}\}dv\right)^2 ds.$$

It will greatly simplify notation to write

$$f_i(v) = (\partial E(u(v))/\partial u_i)[g_i'(u_i(v))]^{-1/2}$$

so that

$$\dot{u}_i = -[g_i'(u_i)]^{-1/2} \int_{t-T}^{t} a_i(t-s) f_i(s) ds$$

and

$$V = 2E(u) - \sum_{i=1}^{n} \int_{t-T}^{t} a_i'(t-s) \Big(\int_s^t f_i(v) dv \Big)^2 ds.$$

Then

$$\begin{aligned} \dot{V} = -2 \sum_{i=1}^{n} \Big\{ & (\partial E / \partial u_i) [g_i'(u_i)]^{-1/2} \int_{t-T}^{t} a_i(t-s) f_i(s) ds \Big\} \\ & + \sum_{i=1}^{n} a_i'(T) \Big(\int_{t-T}^{t} f_i(v) dv \Big)^2 - \sum_{i=1}^{n} \int_{t-T}^{t} a_i''(t-s) \Big(\int_s^t f_i(v) dv \Big)^2 ds \\ & - 2 \sum_{i=1}^{n} \int_{t-T}^{t} a_i'(t-s) \int_s^t f_i(v) dv \, ds f_i(t) \leq 0. \end{aligned}$$

If we integrate the last term by parts we obtain

$$\begin{aligned} -2 \sum_{i=1}^{n} f_i(t) \Big\{ -a_i(t-s) \int_s^t f_i(v) dv \Big|_{s=t-T}^{s=t} - \int_{t-T}^{t} a_i(t-s) f_i(s) ds \Big\} \\ = 2 \sum_{i=1}^{n} f_i(t) \int_{t-T}^{t} a_i(t-s) f_i(s) ds \end{aligned}$$

since $a_i(T) = 0$ by (2); and this cancels with the first term of $\dot{V}$ yielding

$$\dot{V} \leq - \sum_{i=1}^{n} \int_{t-T}^{t} a_i''(t-s) \Big(\int_s^t f_i(v) dv \Big)^2 ds. \tag{10}$$

since $a_i'(t) \leq 0$ by (2). Thus, $E(u(t))$ is bounded using (4*), and so $u(t)$ is bounded.

Let

$$M = \{ u | \partial E(u) / \partial u_i = 0 \text{ for } i = 1, \dots, n \}.$$

Now for differential equations without a delay the invariance principle of LaSalle (cf. Hirsch (1989), for example) would show that this bounded solution approaches M. In this case we have a functional differential equation and the argument of Hale (1977, p.119) will allow us to conclude that the solution approaches M.

4. CONCLUSION

The Cohen-Grossberg-Hopfield model (1) shows that when the synapse has zero threshold, then solutions of (1) converge to the equilibrium points of (1). Our Theorem 2 shows that when the synapse has a threshold, then solutions converge to neighborhoods of those same equilibrium points.

The Cohen-Grossberg-Hopfield model also shows that when the short term average charge in a neuron and at a synapse is approximated by the charge at one point in time, then solutions of (1) converge to the equilibrium points of (1). Theorem 3 shows that when the charge is averaged and when dissipation of the charge is also modeled, then solutions converge to the equilibrium points of (1).

Physical systems have been known to exhibit small oscillations in a neighborhood of the equilibrium points of (1) and there has been much interest in models with a delay that will produce such effects (cf. Marcus and Westerverlt (1989a) and (1989b) for examples, discussion, and extensive bibliographies). These authors introduce pointwise delays, instead of distributed delays with decaying properties, and the resulting solutions of those equations exhibit oscillatory behavior. By contrast our delay model is patterned after the short-term average discussed by Hopfield and no sustained oscillations occur, while they can occur in our threshold model. Thus, we offer these two models as possible evidence that the oscillations observed in physical neural systems may be caused by the threshold instead of the delay. In this sense, it may be more important to have been able to solve (2) and (5) separately, than to solve the combined system (6).

REFERENCES

Burton, T.A. (1985). Stability and Periodic Solutions of Ordinary and Functional Differential Equations. Academic Press: Orlando, Florida.

Cohen, M.A. (1990). The stability of sustained oscillations in symmetric cooperative-competitive networks. Neural Networks, 3, 609–612.

Cohen, M.A. and Grossberg, S. (1983). Absolute stability of global pattern formation and parallel memory storage by competitive neural networks. IEEE Transactions on SMC, SMC 13, 815–826.

Haag, J. (1962). Oscillatory Motions. Wadsworth: Belmont, Calif.

Hale, J.K. (1977). Theory of Functional Differential Equations. Springer: New York.

Hopfield, J.J. (1982). Neural networks and Physical systems with emergent collective computational abilities. Proceedings of the National Academy of Sciences, 79, 2554–2558.

Hopfield, J.J. (1984). Neurons with graded response have collective computational properties like those of two-state neurons. Proceedings of the National Academy of Sciences, 81, 3088–3092.

Levin, J.J. and Nohel, J. (1964). On a nonlinear delay equation. J. Math. Anal. Appl., 8, 31–44.

Marcus, C.M. and Westervelt, R.M. (1989a). Stability of analog neural networks with delay. Physical Rev. A, 39, 347–359.

__________ (1989b). Dynamics of analog neural networks with time delay, in Advances in Neural Information Processing Systems (D.S. Touretzky, ed.). Morgan Kauffman, San Mateo, CA.