Class of Hamiltonian neural networks

Philippe De Wilde

Department of Electrical and Electronic Engineering, Imperial College of Science, Technology and Medicine, London SW7 2BT, United Kingdom

(Received 9 June 1992)

We investigate analog neural networks. They have continuous state variables that depend continuously on time. Although they all have an energy function, not all can have their dynamics derived from a Hamiltonian. Some necessary conditions are given for the network to have Hamiltonian dynamics. We give an example and, using symplectic transformations, describe a whole class of neural networks with Hamiltonian dynamics.

PACS number(s): 87.10.+e, 03.20.+i

I. ANALOG NEURAL NETWORKS AND THE NEED FOR A HAMILTONIAN FORMULATION

Most analog neural networks have the following dynamics:

$$\mu_i \frac{d\,u_i(t)}{dt} = -u_i(t) + \sum_{j=1}^{2n} T_{ij} f_j(u_j(t)) + I_i(t), \quad i = 1, \ldots, 2n, \quad \mu_i > 0. \tag{1}$$

This describes a system with $2n$ neurons, see Fig. 1, and [1, 2]. Neuron i has state $u_i(t)$. This is a continuous function of time t. The *weight* or *synapse* T_{ij} describes the *influence* that neuron j exerts on neuron i. The function f_j is the transfer function of neuron j, it has asymptotes -1 and $+1$, and is monotonically increasing and continuous. There is an external input $I_i(t)$ to neuron i. Each neuron has a positive time constant μ_i.

Most useful results known about neural networks with dynamics (1) are on the existence and stability of equilibria. These results are obtained via the construction of energy functions and Lyapunov functions [3, 2]. An energy function for (1) is

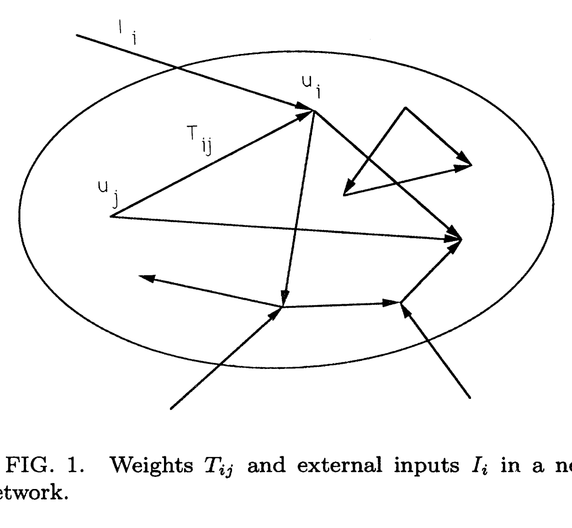

FIG. 1. Weights T_{ij} and external inputs I_i in a neural network.

$$E = -\tfrac{1}{2}\sum_{i=1}^{2n}\sum_{j=1}^{2n} T_{ij} f_i(u_i) f_j(u_j) - \sum_{i=1}^{2n}\left(I_i f_i(u_i) + \int_0^{f_i(u_i)} f_i^{-1}(x)dx \right),$$

with f_i^{-1} the inverse function of f_i. The function E decreases as the network evolves in time (provided the matrix T is symmetric). This shows that all analog neural networks with symmetric connectivity matrix have an energy function. A Lyapunov function for (1) is

$$v = \sum_{i=1}^{2n} \alpha_i u_i^2,$$

with suitable conditions on the α_i [2]. It has a negative definite derivative along the solutions of (1).

This approach provides no information about the time-dependent behavior of the network: the possibility of periodic solutions (oscillatory behavior), phase locking, bifurcations, chaos. However, all these aspects have been studied extensively for Hamiltonian systems [4].

When a Hamiltonian is known for a system, it is also possible to reduce the system (eliminate variables) [5]. Finding a constant of the motion is a first step in the reduction process. The elimination of variables is equivalent to elimination of neurons. This reduces the number of interconnections, and means that Hamiltonian neural networks may have a very-large-scale-integration VLSI layout with a low number of interconnections. This would facilitate an area-efficient VLSI layout, as the interconnections take up most of the space on a neural-network computer chip [6]. We will in fact see at the end of this paper that our class of Hamiltonian networks are economical to rout (connect the neurons in VLSI). This is the original motivation for our research.

The question therefore arises: for which neural networks do Hamiltonians exist? We do not give an exhaustive answer to this question, but will establish an important class of Hamiltonian neural networks. Their dynamics will be slightly different from (1).

II. HAMILTONIAN SYSTEMS

A Hamiltonian system has dynamics

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}, \quad i = 1, \ldots, n. \tag{2}$$

The variables $q_i(t)$ are the generalized coordinates and $p_i(t)$ are the generalized conjugate momenta. The time derivative of $q_i(t)$ is denoted by $\dot{q}_i$. The function $H(q_1, \ldots, q_n, p_1, \ldots, p_n, t)$ is the Hamiltonian. The question we want to solve is how to bring (1) in the form (2). This may not be possible for all systems (1).

In many systems the Hamiltonian is the sum of kinetic and potential energy [7],

$$H = T + V.$$

This occurs, however, mainly in second-order systems, where the moment p_i is proportional to $\dot{q}_i$.

There are two methods of extending (1) to a second-order system (in $\ddot{u}_i$). One is to add a differential equation for the weights [8], e.g.,

$$\tau_{ij}\dot{T}_{ij} = -T_{ij} + S_i(u_i)S_j(u_j), \quad \tau_{ij} > 0$$
$$S_i(u_i) = \frac{1}{1 + e^{-c_i u_i}}, \quad i = 1, \ldots, 2n, \quad j = 1, \ldots, 2n, \tag{3}$$

with c_i constant. This can then be substituted in the derivative of (1). Equation (3) models well biological learning, but the resulting dynamics (in $\ddot{u}_i$) are very difficult to analyze, particularly so because of the different time constants τ_{ij} and μ_i involved in the change of the weights and the change of the neuron states.

Another way is simply to take the derivative of (1). This gives, after substitution of (1),

$$\begin{aligned} \mu_i \ddot{u}_i &= -\dot{u}_i + \sum_j T_{ij} f_j'(u_j)\dot{u}_j + \dot{I}_i \\ &= \frac{1}{\mu_i}\left[u_i - \sum_j T_{ij} f_j(u_j) - I_i + O(u_i u_j)\right] + \dot{I}_i, \\ & \quad i = 1, \ldots, 2n. \end{aligned} \tag{4}$$

We will from now on suppose that the external input is constant in time,

$$\dot{I}_i(t) \equiv 0. \tag{5}$$

This makes (1) an autonomous system.

One can drop second-order interactions between the neurons (products of u_i and u_j). Even with these simplifications the system (4) is very unlike any other second-order dynamical system from mechanics, because of the nonlinear functions f_i. The result of this section is that, if a neural network has a Hamiltonian, it is likely to be as a first order system, with p_i not proportional to $\dot{q}_i$.

III. THE HAMILTONIAN FOR A NEURAL NETWORK WITH BIPARTITE CONNECTIVITY

Few systems of first-order differential equations can be derived from a Hamiltonian via (2). The Dirac equation can be brought in this form [9], but only because the complex conjugate of the wave function has a meaning in quantum mechanics.

The Volterra-Lotka equations [10, 11], modeling predator-prey systems in biology, are more interesting for our purpose. The equations

$$\begin{aligned} \dot{x} &= (a - by)x, \\ \dot{y} &= -(c - fx)y, \quad x, y \geq 0, \quad a, b, c > 0 \end{aligned} \tag{6}$$

can be derived from a Hamiltonian only after the transformation

$$x = e^q, \quad y = e^p. \tag{7}$$

Something very similar will happen for neural networks.

Consider a set of $2n$ neurons. Split them up in two sets of n neurons (see Fig. 2). The graph expressing the connectivity of the neural network is then a bipartite graph [12]. For this reason we will call the network a network with bipartite connectivity. Kosko has shown [8] how powerful a model this is. In his neural network models the neurons are split up in two groups. Each group calculates, alternately, its state by summing the products of interconnection weights T_{ij} and states u_j of neurons in the other group, and then applies a nonlinear function, e.g., f_i. Whittle's antiphon [13] is another example of a network with bipartite connectivity. It is similar to Kosko's model, but the weights are binary and only one group of neurons applies the nonlinear function.

As in the Volterra-Lotka equations, the generalized coordinates will be the state variables of part of the system (here the first n neurons, called the q neurons). The conjugate momenta are the state variables of the other part of the system (the other n neurons, called the p neurons). The state variables of the q neurons will be denoted by $u_i, i = 1, \ldots, n$, those of the p neurons by $w_i, i = 1, \ldots, n$. We will refer to them as q neurons and p neurons, respectively, because the dynamics will be expressed in terms of the variables

$$q_i = f_i(u_i), \quad p_i = g_i(w_i), \quad i = 1, \ldots, n, \tag{8}$$

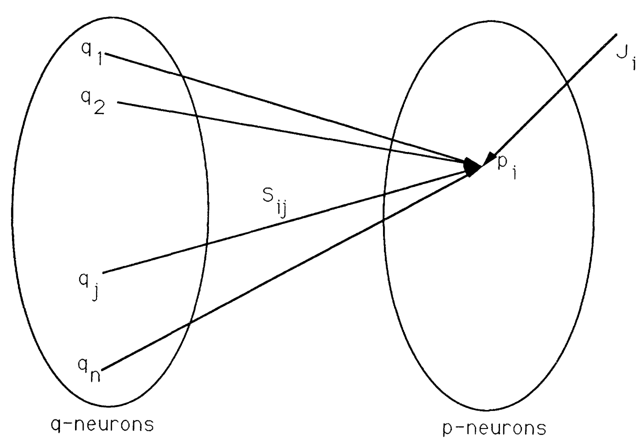

FIG. 2. The network with bipartite connectivity, expressing how the q neurons influence the p neurons.

with all f_i and g_i continuous monotonically increasing functions with asymptotes -1 and $+1$ (for example, $\tanh x$). The reason for choosing this transformation is similar to the reason for choosing (7) for the Volterra-Lotka equations. It can also be remarked that both (7) and (8) are not symplectic. More about transformations of Hamiltonian systems in the last section.

There will be two weight matrices: S and T. The matrix S expresses how the p neurons are influenced by the q neurons, see Fig. 2, and T expresses how the q neurons are influenced by the p neurons, see Fig. 3. The matrices S and T should not be confused with the weight matrix for the complete network of $2n$ neurons. This $2n \times 2n$ matrix has the form

$$\begin{pmatrix} 0 & T \\ S & 0 \end{pmatrix}.$$

External input to the q neurons will be denoted $I_i, i = 1, \dots, n$, to the p neurons $J_i, i = 1, \dots, n$. We choose the input to be constant in time, so that the system is conservative.

The choice of a Hamiltonian is always guesswork if there is no obvious choice for kinetic and potential energy. For this reason a Hamiltonian H is *proposed* now, and its dynamics will be investigated,

$$H = \tfrac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} S_{ij} q_i q_j - \sum_{i=1}^{n} J_i q_i - \tfrac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} T_{ij} p_i p_j + \sum_{i=1}^{n} I_i p_i.$$

The first equation in (2) is now

$$\begin{aligned} \dot{q}_k &= \frac{\partial H}{\partial p_k} \\ &= \frac{\partial}{\partial p_k}\left[-\tfrac{1}{2}\sum_{i(\neq k)} T_{ik} p_i p_k - \tfrac{1}{2}\sum_{j(\neq k)} T_{kj} p_k p_j - \tfrac{1}{2} T_{kk} p_k^2\right] + I_k \\ &= -\tfrac{1}{2}\sum_{i(\neq k)} T_{ik} p_i - \tfrac{1}{2}\sum_{j(\neq k)} T_{kj} p_j - T_{kk} p_k + I_k. \end{aligned}$$

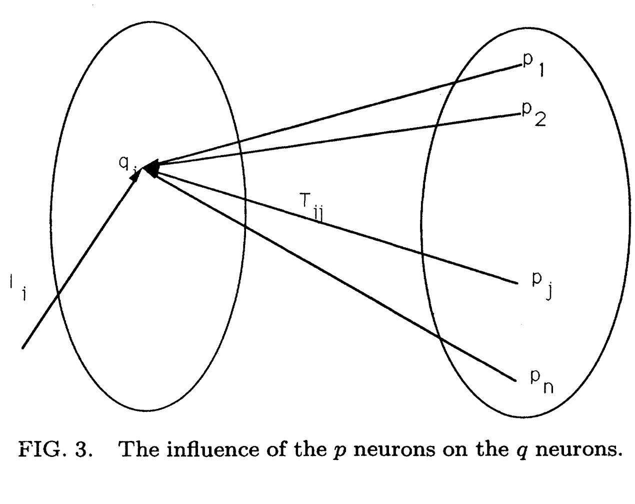

FIG. 3. The influence of the p neurons on the q neurons.

This expression can be simplified by making the assumption that T is symmetric. This will be done here mainly for aesthetic reasons, in order to simplify the expression for $\dot{q}_k$. The assumption that T is symmetric does *not* mean that the synaptic influence between two neurons is symmetric, but that the weight between p_i and q_j is the same as the weight between p_j and q_i. (We remark that we do *not* have to suppose that T has zero diagonal.)

The simplified expression for $\dot{q}$ is now

$$\dot{q}_k = -\sum_{j=1}^{n} T_{kj} p_j + I_k. \tag{9}$$

Similarly, one obtains

$$\begin{aligned} \dot{p}_k &= -\frac{\partial H}{\partial q_k} \\ &= -\sum_{j=1}^{n} S_{kj} q_j + J_k. \end{aligned} \tag{10}$$

It is obvious that for these dynamics

$$\frac{\partial \dot{q}_i}{\partial q_i} = -\frac{\partial \dot{p}_i}{\partial p_i},$$

a necessary condition for Hamiltonian dynamics that can be derived from (2). This condition can never be fulfilled by (1), thus providing another indication that the requirement of bipartite connectivity was necessary to obtain Hamiltonian dynamics for a neural network.

IV. THE NEURONAL DYNAMICS

What dynamics are described by Eqs. (9) and (10)? The equations are similar for q neurons and p neurons, so we will concentrate on (9).

On application of the inverse of the transformation (8), one gets

$$\frac{d}{dt} f_k(u_k) = -\sum_{j=1}^{n} T_{kj} g_j(w_j) + I_k,$$

or

$$\dot{u}_k = \frac{1}{f_k'(u_k)}\left(-\sum_{j=1}^{n} T_{kj} g_j(w_j) + I_k\right).$$

This is Cohen-Grossberg activation dynamics [8] for the q neurons, showing how they are influenced by the states w_j of the p neurons, and an external input I_k to the q neurons.

The non-negative function $1/f_k'(u_k)$ is an amplification function. For example, if $f_k(u_k) = \tanh(u_k)$, then $1/f_k'(u_k) = \cosh^2(u_k)$. This stays bounded, provided that u_k is bounded.

If the initial state of the system

$$\dot{u}_k = \frac{1}{f_k'(u_k)} \left(-\sum_{j=1}^{n} T_{kj} g_j(w_j) + I_k \right),$$
$$\dot{w}_k = \frac{1}{g_k'(w_k)} \left(-\sum_{j=1}^{n} S_{kj} f_j(u_j) + J_k \right),$$
$$k = 1, \ldots, n \tag{11}$$

is sufficiently close to an equilibrium (fixed point), the system will converge to it. The stability of the equilibria of (11) has been studied in adaptive resonance theory [8, 3]. The dynamics show in general no unpredictable, chaotic behavior, but are more like a damped oscillation.

In the usual way [3], it is also possible to change (9) and (10) to

$$\dot{q}_k = w_k - \sum_{j=1}^{n} T_{kj} p_j + I_k,$$
$$\dot{p}_k = u_k - \sum_{j=1}^{n} S_{kj} q_j + J_k,$$

by changing the Hamiltonian to

$$H = \tfrac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} S_{ij} q_i q_j - \sum_{i=1}^{n} \left(J_i q_i - \int_0^{q_i} f_i^{-1}(x) dx \right) - \tfrac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} T_{ij} p_i p_j + \sum_{i=1}^{n} \left(I_i p_i + \int_0^{p_i} g_i^{-1}(x) dx \right)$$

with f_i^{-1} the inverse function of f_i. An additional requirement in this case is $f_i(0) = g_i(0) = 0, i = 1, \ldots, n$.

V. SYMPLECTIC TRANSFORMATIONS AND BIPARTITE CONNECTIVITY

Is it possible to transform the bipartite connectivity network in Figs. 2 and 3 so that the dynamics remain Hamiltonian (2)? This will be the case if the transformation is symplectic [10, 14].

Denote the transformation by

$$h_i = h_i(q_1, \ldots, q_n, p_1, \ldots, p_n), \quad i = 1, \ldots, 2n. \tag{12}$$

Its Jacobian matrix is

$$Dh = \begin{pmatrix} \frac{\partial h_1}{\partial q_1} & \cdots & \frac{\partial h_1}{\partial q_n} & \frac{\partial h_1}{\partial p_1} & \cdots & \frac{\partial h_1}{\partial p_n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_{2n}}{\partial q_1} & \cdots & \frac{\partial h_{2n}}{\partial q_n} & \frac{\partial h_{2n}}{\partial p_1} & \cdots & \frac{\partial h_{2n}}{\partial p_n} \end{pmatrix}.$$

The condition for the transformation h to be symplectic is

$$(Dh)^T \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix} Dh = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}, \tag{13}$$

with I the $n \times n$ identity matrix. Split Dh up into four $n \times n$ block matrices,

$$Dh = \begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

Then the condition (13) becomes

$$\begin{pmatrix} -AC + CA & -AD + CB \\ -BC + DA & -BD + DB \end{pmatrix} = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}.$$

This has several solutions in A, B, C, and D but we will concentrate on two:

$$B = 0, \quad C = 0, \quad D = A^{-1} \tag{14}$$

and

$$A = 0, \quad D = 0, \quad C = B^{-1}. \tag{15}$$

We will consider the following symmetry operations [15] on the network: interchanging two neurons and inverting the sign of a neuron state. We will also suppose that the f_i and g_i are odd functions. We then consider $u_i \leftrightarrow u_j$ or $u_i \leftrightarrow w_j$ or $w_i \leftrightarrow w_j$ or $u_i \leftrightarrow -u_i$ or $w_i \leftrightarrow -w_i$, or, alternatively, $q_i \leftrightarrow q_j$ or $q_i \leftrightarrow p_j$ or $p_i \leftrightarrow p_j$ or $q_i \leftrightarrow -q_i$ or $p_i \leftrightarrow -p_i$. These operations can be composed to form all symmetry operations of the network [15]. Which of them are symplectic?

The weight matrix of the bipartite network is

$$\begin{pmatrix} 0 & T \\ S & 0 \end{pmatrix},$$

where S and T are the $n \times n$ weight matrices from (9) and (10). The places of the zeros in this matrix describe the connectivity. How is it altered by the symplectic transformations?

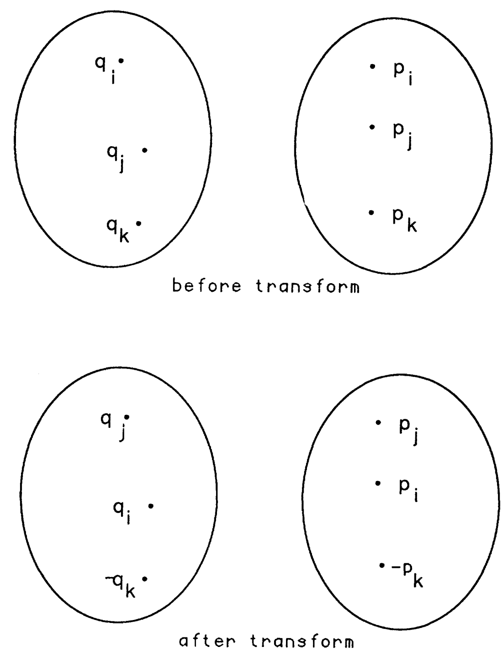

FIG. 4. A neural network before and after a transformation as in (16).

For the symmetry operations of the network (they are linear transformations),

$$h = Dh.$$

We consider now the case where the symmetry transformations are symplectic because they obey (14). The weight matrix is transformed in the following way:

$$\begin{pmatrix} A & 0 \\ 0 & A^{-1} \end{pmatrix} \begin{pmatrix} 0 & T \\ S & 0 \end{pmatrix} = \begin{pmatrix} 0 & AT \\ A^{-1}S & 0 \end{pmatrix}.$$

But the symmetry transformations are their own inverse, so $A = A^{-1}$, and

$$\begin{pmatrix} A & 0 \\ 0 & A^{-1} \end{pmatrix} \begin{pmatrix} 0 & T \\ S & 0 \end{pmatrix} = \begin{pmatrix} 0 & AT \\ AS & 0 \end{pmatrix}. \tag{16}$$

This can be formulated in the following way (see Fig. 4): Any interchange or sign inversion of the states of p neurons, if complemented by the same transformation on the q neurons, still allows the dynamics of the network to be derived from a Hamiltonian.

The transformations that obey (15) transform the weight matrix in the following way:

$$\begin{pmatrix} 0 & B \\ B^{-1} & 0 \end{pmatrix} \begin{pmatrix} 0 & T \\ S & 0 \end{pmatrix} = \begin{pmatrix} BS & 0 \\ 0 & B^{-1}T \end{pmatrix}.$$

Again, the transformations are their own inverse, i.e., $B = B^{-1}$, or

$$\begin{pmatrix} 0 & B \\ B^{-1} & 0 \end{pmatrix} \begin{pmatrix} 0 & T \\ S & 0 \end{pmatrix} = \begin{pmatrix} BS & 0 \\ 0 & BT \end{pmatrix}. \tag{17}$$

This means (see Fig. 5): The connections from q neurons to p neurons are made to run from q neurons to q neurons that correspond to the p neurons. If an interchange or sign inversion is done for the q neurons, and if the same is done for the corresponding p neurons, the dynamics of the network can still be derived from a Hamiltonian.

All the networks with Hamiltonian dynamics that we have considered here are easy to rout for VLSI layout [16], because the connectivity graphs are bipartite, as in (16), or can be split up into two nonconnected parts, as in (17).

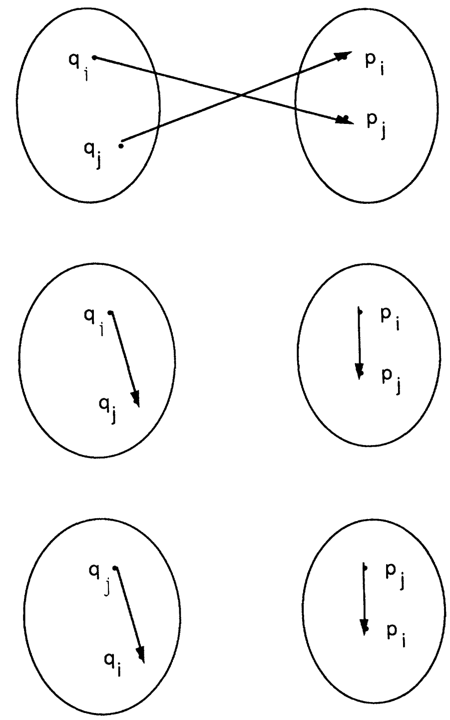

FIG. 5. A transformation as in (17), split up in two steps.

ACKNOWLEDGMENTS

The author would like to thank Dr. D. Arrowsmith, Dr. F. Vivaldi, Professor J. Taylor, and P. Pratt for interesting discussions, and Professor I. Aleksander for his continuing support.

[1] John J. Hopfield and David W. Tank, Science **233**, 625 (1986).
[2] A.N. Michel, J.A. Farrell, and W. Porod, IEEE Trans. Circ. Syst. **CAS-36**, 229 (1989).
[3] S. Grossberg, *The Adaptive Brain* (North-Holland, Amsterdam, 1986), Vols. I and II.
[4] *Hamiltonian Dynamical Systems* (a reprint selection), edited by R. S. MacKay and J. D. Meiss (Adam Hilger, Bristol, 1987).
[5] G. Marmo, E.J. Saletan, A. Simoni, and B. Vitale, *Dynamical Systems* (Wiley, Chichester, 1985).
[6] W.-C. Fang, B.J. Shen, O. T.-C. Chen, and J. Choi, IEEE Trans. Neural Networks **3**, 506 (1992).
[7] L.D. Landau and E.M. Lifshitz, *Mechanics* (Pergamon, Oxford, 1976).
[8] B. Kosko, *Neural Networks and Fuzzy Systems* (Prentice Hall, Englewood Cliffs, NJ, 1992).
[9] P.M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953).
[10] D.K. Arrowsmith and C.M. Place, *An Introduction to Dynamical Systems* (Cambridge University Press, Cambridge, England, 1990).
[11] F. Verhulst, *Nonlinear Differential Equations and Dynamical Systems* (Springer-Verlag, Berlin, 1990).
[12] C. Berge, *Graphs* (North-Holland, Amsterdam, 1985).
[13] P. Whittle., Proc. R. Soc. London Ser. A, **423**, 201 (1989).
[14] R. Abraham and J.E. Marsden, *Foundations of Mechanics* (Benjamin, New York, 1967).
[15] Ph. De Wilde and A.M.C.-L. Ho, in *Artificial Neural Networks II*, Proceedings of the International Conference on Artificial Neural Networks, Amsterdam, 1992, edited by I. Aleksander and J. Taylor (North-Holland, Amsterdam, 1992).
[16] J.D. Ullman, *Computational Aspects of VLSI* (Computer Science, Rockville, MD, 1984).