2005 International Symposium on Nonlinear
Theory and its Applications (NOLTA2005)
Bruges, Belgium, October 18-21, 2005

Influence of Nonlinearity on the Limit-cycles of Neural Networks with Asymmetrical Cyclic Connections

Shinya SUENAGA, Yoshihiro HAYAKAWA, Koji NAKAJIMA

Laboratory for Brainware Systems, Laboratory for Nanoelectronics and Spintronics,
Research Institute of Electrical Communication, Tohoku University
Katahira 2-1-1, Aobaku Sendai 980-8577, Japan,
Email: {shinya, asaka, hello}@nakajima.riec.tohoku.ac.jp

Abstract—We have analysed the effect of nonlinearity of an activation function on the oscillation of the neural network with asymmetrical cyclic connections by using the averaging method. The oscillation obtained by linear approximation, in which the activation function is approximated by a linear function, is multi-modal although the actual oscillation with numerical simulation is unimodal. When the activation function tanh is approximated with a function that includes a nonlinear term, we show that multi-mode oscillations are unstable by using the averaging method.

1. Introduction

An artificial neural network is an information processing system based on simplified models of living nerve-cell networks. It is regarded as the information processing system which has a nature of brain functions, that is, flexibility, distributed processing, parallel computation, and so on. Many problems which are hard to be solved by ordinary methods have been approached by a network with recurrent connections for these systems.

Although most of those applications are used for solving "static" problems, there have recently been many attempts to process temporal information using artificial neural networks as dynamical systems [1]. In these applications, the dynamics of the networks play a very important role. However, understanding the details of a dynamical system is more difficult than that of stable networks, especially in large scale ones.

To analyze the dynamical behaviour of the network comprised of a large number of neural units, we exploit a ring structure. Advantages of the structure is that it removes the effect of a boundary in finite element networks, and that it oscillates easily. Additionally, each unit are connected to each other with different weight values. There are some early works [2, 3] for such network. Both of them reported generation of various limit-cycles. In our previous work [4], we obtained the period of such limit-cycles by using linear approximation of an activation function.

Although the solution obtained by using the linear approximation is the superposition of some frequency components, the solution was different from the limit-cycle with numerical simulation. The oscillation of limit-cycles obtained from linear approximation is multi-modal although the actually oscillation with numerical simulation is unimodal. In this paper, we show by using the averaging method that this difference is caused by the nonlinearity of the activation function.

This paper is organized as follows. In Section 2, the neuron model, the network structure, and a typical limit-cycle of the network are presented. In Section 3, the relation between the period of limit-cycle and various parameters is obtained by using linear approximation. Details of calculations of the averaging method for the limit-cycles are presented in Section 4, and the results are summarized in Section 5.

2. Neural network with asymmetrical cyclic connections

2.1. Model and structure

The limit-cycles considered in this paper are based on conventional neural network equations,

$$\begin{aligned}\tau\frac{du_i(t)}{dt} &= -u_i(t) + \alpha x_i(t) + \sum_{j=0,j\neq i}^{N-1} w_{ij}x_j(t)\\ x(t) &= f(u(t)),\end{aligned} \quad (1)$$

where $x_i(t)$, $u_i(t)$, N, τ, w_{ij}, α, and $f(\cdot)$ are the output and the internal state of unit i, the number of neural units, a time constant, the synaptic weight from unit j to unit i, the self-connection weight ($\alpha > 1$), and a sigmoid activation function, respectively. The sigmoid activation function is given here by

$$f(u) = \tanh(\beta u), \quad (2)$$

where β is a gain parameter ($\beta > 0$).

We use a ring network structure with asymmetrical connections. The network has a cyclic weight matrix $\boldsymbol{W}$. Each unit has synaptic connections from itself and L-neighbor units, where L represents the length of connections and is an integer with $1 \leq L < N/2$. The weight matrix $\boldsymbol{W}$ is

$$
\boldsymbol{W} = \begin{pmatrix} w_0 & w_1 & w_2 & \cdots & w_{N-2} & w_{N-1} \\ w_{N-1} & w_0 & w_1 & \cdots & & w_{N-2} \\ \vdots & & \ddots & & & \vdots \\ w_1 & w_2 & w_3 & \cdots & w_{N-1} & w_0 \end{pmatrix}. \tag{3}
$$

w_{ij} represents the element of $\boldsymbol{W}$ in the i-th row and j-th column. For simplicity, we let $|w_l| = 1$ and $w_l = -w_{N-l}$ for $1 \le l \le L$, namely

$$
\boldsymbol{W} = \begin{cases} w_0 = \alpha \\ w_1 = \cdots = w_L = -1 \\ w_{N-1} = \cdots = w_{N-L} = +1 \\ \text{otherwise} = 0 \end{cases}. \tag{4}
$$

2.2. Limit-cycles

These neural network exhibits a lot of limit-cycles with the characteristics which transmit the information about the state of the unit in one direction in the ring network. Let the unit i be represented by a black circle if $x_i \ge 0$ and a white circle if $x_i < 0$. Figure 1 shows all spatial patterns of the limit-cycle with $N = 10$ and $L = 1$. Only one circle at each boundary between a white and a black domains changes to opposite color, such that each domain moves in the same direction. Hence, the spatial patterns in Fig. 1 actually seem to rotate. The waveform of the internal state for two units in the limit-cycle is shown in Fig. 2. The waveforms for other units have the same appearance but with a phase shift. Such characteristics may be figured as

$$
x_i = f(u_i) = f\left\{A_i \sin\left(2\pi\frac{K}{N}i - \omega t\right)\right\}, \tag{5}
$$

where A_i, K and ω are an amplitude, a wavenumber, and an angular frequency, respectively. Therefore, we refer to the limit-cycle mentioned above as a sine type limit-cycle. The following relation has been obtained [3],

$$
M = \lfloor (N-1)/2 \rfloor \qquad (L = 1), \tag{6}
$$

where M is the number of limit-cycles and $\lfloor h \rfloor$ indicates the greatest integer smaller than or equal to h.

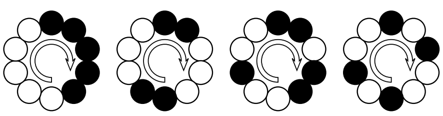

Figure 1: All spatial patterns of the limit-cycle at a certain instant in the neural network with asymmetrical cyclic connections for $N = 10$ and $L = 1$. Black circles and white circles represent the units whose outputs are positive and negative, respectively. For Eq. (5), these patterns are assigned with $K = 1, 2, 3, 4$ sequentially from the left hand diagram.

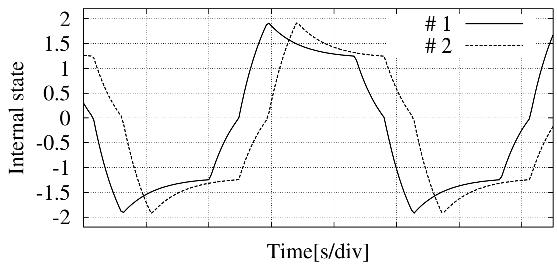

Figure 2: Time evolution of the internal state in a limit-cycle for $N = 10$, $L = 1$, and $K = 1$. All the units show the same internal state except for the phase shift.

3. Linear approximation

The relation between the period of a sine type limit-cycle and the parameters N, K, L, and α are obtained by using linear approximation. The period dependence for these parameters is nonlinear. For example, the period is inverse proportional to L as shown in Fig. 3. To study the neural network with asymmetrical cyclic connections, we need to use a numerical simulation for the activation function $\tanh(\beta u)$. However, $\tanh(\beta u)$ can be approximated by simple functions in the region $\beta \sim 1/\alpha$. A linear function can be used in this region.

The Taylor expansion of tanh is written as

$$
x = \tanh(\beta u) = \beta u - \frac{2\beta^3}{3!}u^3 + \frac{16\beta^5}{5!}u^5 + \cdots. \tag{7}
$$

Thus, in the vicinity $u = 0$, the following linear approximation holds:

$$
x = \beta u. \tag{8}
$$

With this approximation, we can use the method of a linear dynamical system.

3.1. State equation in vector form

With linear approximation, the state equation Eq. (8) can be recast in vector form as

$$
\frac{d}{dt}\boldsymbol{x} = \frac{1}{\tau}\left(\beta \boldsymbol{W}\boldsymbol{x} - \boldsymbol{x}\right) = \boldsymbol{A}\boldsymbol{x}, \tag{9}
$$

where $\boldsymbol{x} = (x_1, x_2, \ldots, x_N)^T$ is the vector of output, and $\boldsymbol{W}$ is given by Eq. (3). Therefore, for example, $\boldsymbol{A}$ with $L = 1$ is written as

$$
\boldsymbol{A} = \frac{1}{\tau}\begin{pmatrix} \beta\alpha - 1 & -\beta & 0 & \cdots & 0 & \beta \\ \beta & \beta\alpha - 1 & -\beta & 0 & \cdots & 0 \\ \vdots & & \ddots & & & \vdots \\ -\beta & 0 & \cdots & 0 & \beta & \beta\alpha - 1 \end{pmatrix}. \tag{10}
$$

$\boldsymbol{x}^T$ is the transpose of $\boldsymbol{x}$. Because $\boldsymbol{A}$ is a $N \times N$ circulant matrix, it can be diagonalized easily [5],

$$
\boldsymbol{Z}^* \boldsymbol{A} \boldsymbol{Z} = \frac{1}{\tau}\operatorname{diag}(\lambda_0, \lambda_1, \ldots, \lambda_{N-1}), \tag{11}
$$

where eigenvalue λ_i $(i = 0, 1, \ldots, N-1)$ is written as

$$\lambda_i = \left[\beta\alpha - 1 - 2\beta\,\mathrm{i}\sum_{j=1}^{L} \sin\left(\frac{2\pi i\, j}{N}\right)\right], \qquad (12)$$

the unitary matrix $\boldsymbol{Z} = \{z_{ij}\}$ $(i, j = 1, 2, \cdots, N)$ is defined by

$$z_{ij} = \frac{1}{\sqrt{N}}\zeta^{(i-1)(j-1)}, \quad \zeta = \exp\left(\frac{2\pi}{N}\,\mathrm{i}\right), \qquad (13)$$

and $\boldsymbol{Z}^*$ is the conjugate transpose matrix of $\boldsymbol{Z}$.

3.2. Solution of the linear approximation

For the initial condition $\boldsymbol{x}(0) = \boldsymbol{P} \in \boldsymbol{R}^N$, we can obtain the following unique solution [6]:

$$\begin{aligned} x_i(t) \;=\; & \frac{1}{\sqrt{N}} e^{\frac{1}{\tau}(\beta\alpha-1)t} \Bigg\{ P_0' + (-1)^{i-1} P_{N/2}' \\ & + 2\sum_{s=1}^{\lfloor (N-1)/2 \rfloor} |P_s'| \sin\left[\left(\frac{2\pi(i-1)s}{N} - \omega_s t\right) + \phi_s\right]\Bigg\} \\ & \Bigg(\text{If } N \text{ is odd, } P_{N/2} = 0.\ |P_s'| = \sqrt{X_{P_s'}^2 + Y_{P_s'}^2}, \\ & \sin\phi_s = \frac{X_{P_s'}}{|P_s'|}, \cos\phi_s = -\frac{Y_{P_s'}}{|P_s'|}\Bigg), \end{aligned} \qquad (14)$$

where $\boldsymbol{P'} = \boldsymbol{Z}^*\boldsymbol{P}$, $P_j' = X_{P_j'} + \mathrm{i}\,Y_{P_j'}$ for $0 < j < N/2$, the phase $\theta_i^s = 2\pi\,(i-1)\,s/N$, and the angular frequency is defined by

$$\omega_s = \frac{2}{\tau\alpha}\sum_{j=1}^{L}\sin\left(\frac{2\pi\, s\, j}{N}\right). \qquad (15)$$

The solution obtained by using linear approximation is the superposition of $M = \lfloor (N-1)/2 \rfloor$ frequency components and some constant components.

Each frequency component of Eq. (14) is extremely similar to a sine type limit-cycle. However, the oscillation obtained from the linear approximation is multi-modal although the actually oscillation with numerical simulation is unimodal like Eq. (5). This is because the activation function does not have nonlinearity in linear approximation. As reported in a next section, if the activation function $\tanh(\beta u)$ is approximated by the function that has a nonlinear term, multi-mode oscillations are unstable.

4. Averaging method

We can analyze the effect of nonlinear characteristics in Eq. (1) by using the method of averaging [7]. If the activation function $\tanh(\beta u)$ is approximated by the function up to the second term of Eq. (7), this function includes a nonlinear term. The first two terms of the Taylor expansion is

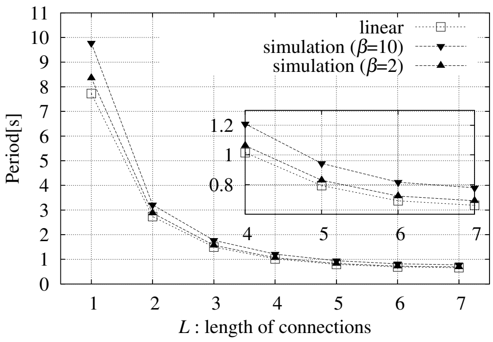

Figure 3: Period of sine type limit-cycles as a function of L obtained by using linear approximation and simulation for $N = 15$, $K = 1$, $\alpha = 1$ and $\tau = 1$.

$$x = f(x) = \beta u - \frac{2\beta^3}{3!}u^3. \qquad (16)$$

Substituting Eq. (16) into Eq. (1), we have

$$\frac{d}{dt}\boldsymbol{u} = -\boldsymbol{u} + \boldsymbol{W}\left(\beta\boldsymbol{u} - \frac{\beta^3}{3}\boldsymbol{u}_c\right) = \boldsymbol{A}\boldsymbol{u} - \frac{\beta^3}{3}\boldsymbol{W}\boldsymbol{u}_c, \qquad (17)$$

where $\boldsymbol{u}_c = \left(u_1{}^3, u_2{}^3, \ldots, u_N{}^3\right)^T$.

Equation (17) reduces to

$$\dot{\boldsymbol{v}} = \boldsymbol{B'}\boldsymbol{v} - \epsilon\boldsymbol{g}, \qquad (18)$$

where $\boldsymbol{B'} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$, $\boldsymbol{u} = \boldsymbol{P}\boldsymbol{v}$, $\epsilon = \beta^3/3$, and $\boldsymbol{g} = \boldsymbol{P}^{-1}\boldsymbol{W}\boldsymbol{u}_c$. The matrix $\boldsymbol{B'}$ is

$$\begin{cases} b'_{i\,i} = \alpha\beta - 1 & i = 1, \ldots, N \\ b'_{2j\,2j+1} = -\omega_j & j = 1, \ldots, M \\ b'_{2j+1\,2j} = \omega_j & j = 1, \ldots, M. \end{cases}, \qquad (19)$$

where

$$\omega_j = 2\beta\sum_{l=1}^{L}\sin\left(\frac{2\pi j\, l}{N}\right) \qquad (20)$$

and the matrix $\boldsymbol{P}$ is defined by

$$\begin{cases} p_{i\,1} = \sqrt{\dfrac{1}{N}},\ p_{i\,N} = (-1)^{i-1}\sqrt{\dfrac{1}{N}}\ (\text{if } N \text{ is even}) \\ p_{i\,2j} = -\sqrt{\dfrac{2}{N}}\sin\dfrac{2\pi j\,(i-1)}{N} \\ p_{i\,2j+1} = \sqrt{\dfrac{2}{N}}\cos\dfrac{2\pi j\,(i-1)}{N} \\ \qquad i = 1, \ldots, N, j = 1, \ldots, M. \end{cases}, \qquad (21)$$

where $\boldsymbol{P}^{-1} = \boldsymbol{P}^T$ holds.

We suppose that $\alpha\beta - 1$ is as small as ϵ in Eq. (18). Involving elements in the principal diagonal of $\boldsymbol{B'}$ in the term of ϵ, Eq. (18) is reduced to

$$\dot{\boldsymbol{v}} = \boldsymbol{D}\boldsymbol{v} - \epsilon\boldsymbol{f}, \qquad (22)$$

where $\boldsymbol{D} = \boldsymbol{A} - (\alpha\beta - 1)\,\boldsymbol{E}$, $\boldsymbol{f} = \boldsymbol{g} - \frac{\alpha\beta-1}{\epsilon}\boldsymbol{E}\boldsymbol{v}$, $\boldsymbol{E}$ is the unit matrix.

With the condition that ϵ is small enough to assume that the nonlinearity of the system is weak, the averaging method can be applied to Eq. (22). If $\epsilon = 0$, the solutions of the equation are

$$\begin{cases} v_1 = \rho_0 \\ v_{2j} = \rho_j \cos\psi_j, \quad j = 1, 2, \ldots, M \\ v_{2j+1} = \rho_j \sin\psi_j, \quad j = 1, 2, \ldots, M \\ v_{\frac{N}{2}} = \rho_{\frac{N}{2}} \quad \text{(if } N \text{ is even)} \end{cases}, \tag{23}$$

where $\psi_j = \omega_j t + \theta_j$. ρ_j and θ_j are constant values determined by an initial condition.

On the basis of Eq. (23) and the averaging method, the autonomous averaged equations are obtained by

$$\begin{aligned} \dot{\rho}_0 &= -\epsilon \langle f_1 \rangle \\ \dot{\rho}_j &= -\epsilon \langle f_{2j+1} \sin\psi_j + f_{2j} \cos\psi_j \rangle \\ \dot{\rho}_{\frac{N}{2}} &= -\epsilon \langle f_N \rangle \quad \text{(if } N \text{ is even)} \\ \dot{\theta}_j &= -\epsilon \langle (f_{2j+1} \cos\psi_j - f_{2j} \sin\psi_j) / \rho_j \rangle , \end{aligned}$$

where $\langle \ \rangle$ indicates

$$\langle f(t) \rangle = \lim_{T\to\infty} \frac{1}{T} \int_0^T f(t) dt. \tag{24}$$

Equation (23) describes the behavior of the original system approximately. Calculating $\boldsymbol{f}$, one obtains

$$\begin{cases} \dot{\rho}_0 = -\epsilon \dfrac{\alpha\rho_0}{N} \left({\rho_0}^2 + 3 \sum\limits_{l=1}^{N/2} {\rho_l}^2 - \dfrac{\alpha\beta - 1}{\alpha} \dfrac{N}{\epsilon} \right) \\ \dot{\rho}_j = -\epsilon \dfrac{\alpha\rho_j}{N} \left(\dfrac{3}{2} {\rho_j}^2 + 3 \sum\limits_{\substack{l=0 \\ l\neq j}}^{N/2} {\rho_l}^2 - \dfrac{\alpha\beta - 1}{\alpha} \dfrac{N}{\epsilon} \right) \\ \dot{\rho}_{\frac{N}{2}} = -\epsilon \dfrac{\alpha\rho_{\frac{N}{2}}}{N} \left(\rho_{\frac{N}{2}}^2 + 3 \sum\limits_{l=0}^{N/2-1} {\rho_l}^2 - \dfrac{\alpha\beta - 1}{\alpha} \dfrac{N}{\epsilon} \right) \\ \qquad \qquad \text{(if } N \text{ is even)} \end{cases} \tag{25}$$

$$\dot{\theta}_j = -\epsilon \frac{\omega_j}{\beta N} \left(\frac{3}{2} {\rho_j}^2 + 3 \sum_{l=0, l\neq j}^{M} {\rho_l}^2 \right) \tag{26}$$

The stability of a stationary states of oscillatory modes are determined by eigenvalues of the corresponding Jacobian. Estimating Jacobian for many fixed points of Eq. (25) numerically, we could not find stable multimode oscillations.

Stable amplitudes of unimodal oscillations are

$$\begin{cases} \rho_0 = \sqrt{\dfrac{N}{\epsilon} \dfrac{\alpha\beta - 1}{\alpha}} \\ \rho_j = \sqrt{\dfrac{2N}{3\epsilon} \dfrac{\alpha\beta - 1}{\alpha}} \quad j = 1, 2, \ldots, M \\ \rho_{\frac{N}{2}} = \sqrt{\dfrac{N}{\epsilon} \dfrac{\alpha\beta - 1}{\alpha}} \quad \text{(if } N \text{ is even)} \end{cases}, \tag{27}$$

and corresponding frequency is $\omega_j / \alpha\beta$.

5. Conclusion

In this paper, we have analysed the effect of nonlinearity of the activation function on the oscillation of the network by using the averaging method. When the activation function $\tanh(\beta u)$ has been approximated by a linear function, i.e. the first term of Taylor expansion, the solution of oscillation has been the superposition of some frequency components. However, the actually oscillation with numerical simulation has been unimodal like Eq. (5). When the activation function $\tanh(\beta u)$ has included a nonlinear term, i.e. the second term of Taylor expansion, we show that multimode oscillations are unstable by using the averaging method. Therefore, we think that the dynamics affected by nonlinear characteristics extract the most major frequency from various components in the solution, obtained by linear approximation, Eq. (14).

Acknowledgments

The authors would like to thank Prof. T. Endo of Meiji University for his valuable advice on the averaging method.

References

[1] J. F. Kolen and S. C. Kremer: "A field guide to dynamical recurrent networks", IEEE Press, New York (2001).

[2] C. Y. Park, Y. Hayakawa, K. Nakajima and Y. Sawada: "Limit cycle of one-dimensional neural networks with the cyclic connection matrix", IEICE Transactions on Fundamentals of electronics, communications and computer sciences, **E79-A**, 6, pp. 752–757 (1996).

[3] G. Setti, P. Thiran and C. Serpico: "An approach to information propagation in 1-D Cellular Neural Networks -Part II: Global propagation", IEEE Transactions on Circuit and Systems I: Fundamental theory and applications, **45**, 8, pp. 790–811 (1998).

[4] S. Suenaga, Y. Hayakawa and K. Nakajima: "Analysis of limit-cycles on neural networks with asymmetrical cyclic connections using approximately activation functions", Lecture Notes in Computer Science, **3213**, pp. 974–980 (2004).

[5] P. J. Davis: "Circulant Matrices", Wiley-Intercsience, New York (1979).

[6] M. W. Hirsch and S. Smale: "Differential equations, dynamical systems, and linear algebra", Academic Press, New York (1974).

[7] J. K. Hale: "Ordinary Differential equations", Robert E. Krieger Pub. Co., Malabar, Fla. (1980). address may be Florida.