TR－H－012

Oscillatory Neural Network and Learning of Continuously Transformed Patterns

Yukio Hayashi

1993. 7. 16

ATR人間情報通信研究所

〒619-02 京都府相楽郡精華町光台2-2　☎07749-5-1011

ATR Human Information Processing Research Laboratories

2-2, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-02 Japan

Telephone: +81-7749-5-1011
Facsimile: +81-7749-5-1008

Oscillatory Neural Network and Learning of Continuously Transformed Patterns *

Yukio Hayashi

ATR Human Information Processing Research Laboratories

2-2, Hikaridai
Seika-cho, Soraku-gun
Kyoto 619-02, Japan

Tel(Direct) +81-7749-5-1073

July 15, 1993

Abstract

We investigate experimentally the dynamic behaviors of an oscillatory neural network. Computer simulations show an interesting characteristic: the autonomous generation of a limit cycle near a memory (memory retrieval with ambiguous fluctuation) for an input near a memory, and of a chaotic orbit among memories (autonomous search) for an input far from memories. We also analyze theoretically a few restricted behaviors near a memory.
This type of neural network can treat spatiotemporal pattern processing in the brain. As an example of dynamic information processing, it is shown that continuously transformed pattern cycles for three Japanease characters are embedded in the limit cycles of the oscillatory neural network by a learning method. The existence in the brain of a continuously transformed pattern operation for a character is also discussed from the cognitive psychological point of view.
The characteristic behavior of limit cycle or chaos according to an input in our oscillatory neural network may be useful for developing a dynamic information processing mechanism for a spatiotemporal pattern in the brain.

Keywords- Oscillatory neural network, Dynamic information processing, Spatiotemporal pattern, Limit cycle near a memory, Chaotic orbit among memories, Continuous transformation, Learning of recurrent network

*This paper is submitted to Journal of Neural Networks, Pergamon Press.

1 INTRODUCTION

This paper investigates the dynamic behaviors of an oscillatory neural network, mainly through computer simulations. We show an interesting characteristic: the autonomous generation of a limit cycle for an input near a memory, and of a chaotic orbit among memories for an input far from memories. We also analyze theoretically a few restricted behaviors near a memory.
This type of neural network can treat spatiotemporal pattern processing in the brain. As an example, it is shown that continuously transformed character patterns are embedded in the limit cycles of the oscillatory neural network by a learning method.

As a motivation of this paper from physiology, oscillatory phenomena have recently been observed in the visual cortex and olfactory bulb in mammalian brains. In the cat visual cortex (Areas 17 and 18), stimulus-specific synchronized oscillations of 40-60 Hz have been reported by both Eckhorn et al. (1988) and Gray et al. (1989, 1990) using moving bar stimuli.
Eckhorn et al. recorded coherent or synchronized oscillations between orientation cells of area 17 and movement direction cells of area 18, at distant cortical positions within a cortical column. And Gray et al. found that there are synchronized oscillations only between cells with similar orientation preferences, and which are separated by large cortical distances.
These synchronized oscillations have received much attention as a feature-binding mechanism in the brain (Eckhorn et al., 1990b; Lummer & Huberman, 1992; Sompolinsky et al., 1991).
Freeman et al. (Sakarda & Freeman, 1987; Yao & Freeman, 1990) have reported that, in rabbit olfactory bulb, near-limit cycle activity occurred for a perceptible specific odor and also that chaotic activity occurred for a novel odor. However, the chaotic dynamics is not considered a disordered randomness but a new type of information processing with an autonomous search mechanism (Tsuda et al., 1987; Mori, Davis & Nara, 1989; Yao & Freeman, 1990).
Thus, oscillatory neural networks are of interest not only in physiology, but also in dynamic information processing for a spatiotemporal pattern in the brain.

To generate a spatiotemporal pattern or time-sequential pattern, an asymmetrically connected network, e.g., excitatory and inhibitory (E-I) connections, is important (Amari, 1972b). In the state of the art, several types of artificial or physiological oscillatory neural network models have been reported. These networks can be classified roughly into the four following types.

1. A network with mutual E-I cell connections (Baird, 1986; Lie & Hopfield, 1989; Yao & Freeman, 1990; Grossberg & Somers, 1991).

2. An associative memory model with asymmetric random connections (Mori, Davis & Nara, 1989; Wang, Picheler & Ross, 1990; Renals & Rohwer, 1990).

3. A network based on the E-I effect with dynamical thresholds or time delays as an adaptation (Matsuoka, 1987; Eckhorn et al., 1990a; Wang et al., 1990; Horn & Usher, 1991; Kőnig & Schillen, 1991).

4. Nonlinear coupled oscillators (Sompolinsky et al., 1991; Niebur et al., 1991; Kuramoto, 1991; Lummer & Huberman, 1992).

Since these network models are complex nonlinear systems, the theoretical analyses in the general case are intractable. However, early research by Amari (1971, 1972a) pointed out the oscillations by the mutual effect of E-I cells, and analyzed theoretically an oscillatory condition. In the case of a discrete state (threshold element networks), the theoretical analyses of the stability of an equilibrium state and state transition are reported (Amari, 1972b). From the above suggestions, we consider an oscillatory neural network with E-I pairs.

On the other hand, in order to embed a spatiotemporal pattern into an orbit of a network, learning is necessary. For continuous spatiotemporal patterns, a learning method has been proposed independently by Sato (1990) and Pearlmutter (1989). Sato et al. have shown that a Lorentz attractor as low-dimensional chaos (Sato et al., 1990a), and fluctuations of voice waveforms (Sato et al., 1990b), can be obtained by using the learning algorithm for a recurrent network.
However, it is an issue whether the computation for the learning algorithm becomes very large because the bifurcations cause structural instability. In a recurrent neural network, the learning for a given case is intractable because of bifurcations (Doya, 1992).

In an attempt to avoid the above problem, we improved on Sato's learning method for the oscillatory neural network, which is a specialized recurrent neural network. For hand-written Japanese hiragana characters, the continuously transformed pattern cycles are embedded in the limit cycle of the oscillatory neural network by the improved learning method without the temporal inverse process.
Also, we discuss a continuous transformation of perception from the cognitive psychological point of view. The existence in the brain of such a continuously transformed pattern operation for a character is suggested.

The characteristic behavior of a limit cycle or chaos according to an input in our oscillatory neural network may be useful for developing a dynamic information processing mechanism for a spatiotemporal pattern in the brain.

2 OSCILLATORY NEURAL NETWORK

It is plausible that oscillations in the brain are basically due to the mutual effects of E-I cells. In this paper, we consider the following oscillatory neural network with a simple network architecture among E-I cells, which has an interesting characteristic obtained from the later simulation results (in section 4 and 5).
The network is constructed of excitatory neuronal groups and inhibitory neuronal groups, both of which include N-cells. In the network, each cell is considered not a single neuron but a mean-field approximation of neuronal groups (Wilson & Cowan, 1973; Schuster & Wagner, 1990).
Each excitatory cell is connected to an inhibitory cell to form a corresponding pair and has mutual connections to other excitatory cells. On the other hand, each inhibitory cell has only one connection to the corresponding excitatory cell. The network architecture is shown in Fig. 1. The dynamic equations for their activation are as follows.

$$\dot{x}_i = -x_i + G\left(\sum_{j=1}^{N} W_{ij}x_j - K_{EI}^{(i)}y_i + I_i\right), \tag{1}$$

$$\dot{y}_i = -y_i + G\left(K_{IE}^{(i)}x_i\right), \tag{2}$$

$$G(z) = \frac{2}{\pi}\arctan\left(\frac{z}{a}\right), \tag{3}$$

$$\begin{cases} W_{ij} = \frac{1}{N}\sum_{\alpha=1}^{M}\xi_i^\alpha\xi_j^\alpha & (N=M), \\ W_{ij} = \frac{1}{N}\sum_{\alpha=1}^{M}\xi_i^\alpha\xi_j^\alpha + \delta_{ij} & (N>M). \end{cases} \tag{4}$$

Where $\dot{x}_i$ denotes the time-differential of the variable x_i, each x_i and y_i are the averaged pulse density for excitatory and inhibitory cells, respectively, $G(z)$ is a sigmoid function from averaged

membrane potential to averaged pulse density and a is the slope of $G(z)$.
The mutual connection W_{ij} between the ij-th excitatory cells is defined by the correlation matrix of M's memory pattern ξ^α (each element is a binary value, ± 1), and δ_{ij} denotes Kronecker's delta. From the characteristics of the correlation matrix, it is clear that $W_{ii} = 1$ *or* $(N+M)/N$ ($N = M$ *or* $N > M$) and $W_{ij} = W_{ji}$. The connections for the i-th E-I pair are inhibitory: $-K_{EI}^{(i)}$ and excitatory: $K_{IE}^{(i)}$. I_i is the input or input bias for the i-th excitatory cell. The output pattern is given by the activation values of excitatory cells; $x_i(t)$.
In the next section, in order to simplify the stability analyses near a memory in the state space, we consider an oscillatory neural network with homogeneous E-I pairs, which have the same connection weight values K_{IE} and K_{EI}. However, when the network is learned (in section 5), the parameter $K_{EI}^{(i)}$ is changeable for the cell number i.

3 RESTRICTED STABILITY ANALYSES

Theoretical analysis of the oscillatory neural network is intractable in the general case. To investigate a typical behavior of the oscillatory neural network, we ran computer simulations, which results are shown in the next section. However, we can theoretically analyze a few restricted behaviors near a memory in the state space, which arise in simulation.

3.1 Perturbation Analysis of an Asymptotical Stable Point

In the network without coupling pairs ($K_{IE} = 0$) which is considered to be an associative memory model with an input bias; $\dot{x}_i = -x_i + G(\sum_{j=1}^{N} W_{ij}x_j + I_i) \equiv F_i(x_1, x_2, \ldots, x_N)$, an asymptotical stable point near a memory ξ^α is assumed for an input bias, i.e., there are negative real parts of the eigenvalues for the linearized equations at the equilibrium point $\bar{x}_i \approx \pm 1$. Since an input bias similar to a memory pattern tends to stabilize the memory pattern, the assumption is expected for a large input bias similar to a memory pattern. The analysis of a condition for the assumptions is beyond the scope of this paper.

Under the assumption, when the value of K_{IE} is sufficiently small, we show the existence of a similar asymptotical stable point in the oscillatory neural network by perturbation analysis.

The variable y_i is converted to the following term on the nullcline $\dot{y}_i = 0$, that is $y_i = G(K_{IE}x_i)$. For a sufficiently small value $K_{IE} \equiv \varepsilon > 0$,

$$y_i = G(K_{IE}x_i) = G(0) + \varepsilon x_i G'(0) + O(\varepsilon^2) = \frac{2x_i}{\pi a}\varepsilon + O(\varepsilon^2). \tag{5}$$

Here $G'(z) = 2a/\pi(z^2 + a^2)$ denotes the derivative of $G(z)$.
On the nullcline: $y_i \approx \frac{2x_i}{\pi a}\varepsilon$, the equation (1) is approximated to

$$\dot{x}_i = -x_i + G(\sum_{j=1}^{N} W_{ij}x_j - \frac{2K_{EI}\varepsilon}{\pi a}x_i + I_i) \equiv \tilde{F}_i(x_1, x_2, \ldots, x_N). \tag{6}$$

Note that the self-excitatory connection weight W_{ii} is slightly changed: $W'_{ii} = W_{ii} - \frac{2K_{EI}\varepsilon}{\pi a} + O(\varepsilon^2)$.

We consider the solutions, i.e., equilibrium points, of $\tilde{F}_i(x_1, x_2, \ldots, x_N) = 0$.
At a solution $\{\tilde{x}_i\}$, we obtain the following matrix-vector form of perturbation expansion,

$$[\tilde{F}_i]_{x_k=\tilde{x}_k} = [F_i]_{x_k=\bar{x}_k} + [\partial F_i/\partial x_j]_{x_k=\bar{x}_k}\,\delta x + diag[\partial F_i/\partial W_{ii}]\,\delta W_{ii} + (higher\ order). \tag{7}$$

By the assumption, the first term $[F_i]_{x_k=\bar{x}_k}$ is zero.
Assuming that the Jacobian matrix is invertible, there will be a fixed point at $\bar{x} + \delta x$ where $\delta x = -\,[\partial F_i/\partial x_j]^{-1}_{x_k=\bar{x}_k}\; diag[\partial F_i/\partial W_{ii}]_{x_k=\bar{x}_k}\; \delta W_{ii} \sim O(\varepsilon)$. Therefore, a solution of $[\tilde{F}_i] = 0$ exists, which is a slightly changed equilibrium point: $\tilde{x}_i = \bar{x}_i + \varepsilon x'_i + O(\varepsilon^2)$.

Next, we discuss the stability of the slightly changed equilibrium point $\{\tilde{x}_i\}$.
In the oscillatory neural network, the Jacobian matrix of the linearized equation at the equilibrium point $\{\tilde{x}_i\}$ is,

$$\begin{bmatrix} -1 + W'_{ii}G'_{e1} & W_{12}G'_{e1} & \cdots & W_{1N}G'_{e1} \\ W_{21}G'_{e2} & -1 + W'_{ii}G'_{e2} & \cdots & W_{2N}G'_{e2} \\ \vdots & \vdots & \vdots & \vdots \\ W_{N1}G'_{eN} & W_{N2}G'_{eN} & \cdots & -1 + W'_{ii}G'_{eN} \end{bmatrix}. \tag{8}$$

Where the i-th function value G'_{ei} is $G'(W'_{ii}\tilde{x}_i + \sum_{j\neq i} W_{ij}\tilde{x}_j + I_i)$.
By a simple calculation of

$$G'_{ei} = G'(\sum_{j=1}^{N} W_{ij}\bar{x}_j + I_i) + \varepsilon(\sum_{j=1}^{N} W_{ij}x'_j - \frac{2K_{EI}\bar{x}_i}{\pi a})\; G''(\sum_{j=1}^{N} W_{ij}\bar{x}_j + I_i) + O(\varepsilon^2), \tag{9}$$

the Jacobian matrix has the form $A + \varepsilon B + O(\varepsilon^2)$. The matrix A is the Jacobian matrix of the non-coupling system at the equilibrium point $\{\bar{x}_i\}$, whose eigenvalues have negative real parts by the assumption. And the matrix εB is a perturbation matrix.

Note that the equilibrium point is asymptotically stable for a sufficiently small $K_{IE} \equiv \varepsilon$, because the matrix B is bounded by $0 < G''(z) \leq 1/\pi a^2$, $-1 < \bar{x}_i, \tilde{x}_i < 1$, $W_{ii} = 1$ *or* $(N+M)/N$, $W_{ij} = O(1/N)$. Thus, we obtained the following result.

Theorem 1 *It is assumed that there is an asymptotical stable point near a memory in the non-coupling system, and that the Jacobian matrix at the equilibrium point is invertible.*
Under the assumptions, the oscillatory neural network with a sufficiently small coupling parameter $K_{IE} \equiv \varepsilon$ also has a similar asymptotical stable point.

3.2 The Necessary Condition of the Oscillation for one E-I pair

We discuss the necessary condition of the Hopf bifurcation for one E-I pair. If only an E-I pair is destabilized under the other stable pairs at an equilibrium point near a memory, the condition can be applied for the case of the oscillatory neural network with E-I pairs. Though this is not a general case but a special case to produce a oscillation in this network, we encounter it in the computer simulation.

In the following, the discussion is restricted in one E-I pair. It is assumed that there is only one equilibrium point $(\bar{x}_i, \bar{y}_i)$ for a large input bias $I_i > 0$ (see Fig. 2).
The linearized equation at the equilibrium point $(\bar{x}_i, \bar{y}_i)$ is,

$$\begin{pmatrix} \dot{\delta x} \\ \dot{\delta y} \end{pmatrix} = \begin{pmatrix} -1 + W_{ii}G'(v_e) & -K_{EI}G'(v_e) \\ K_{IE}G'(v_i) & -1 \end{pmatrix} \begin{pmatrix} \delta x \\ \delta y \end{pmatrix}. \tag{10}$$

Here v_e, v_i are $v_e \equiv W_{ii}\bar{x}_i - K_{EI}\bar{y}_i + I_i, v_i \equiv K_{IE}\bar{x}_i$, respectively.
Since the eigenvalues λ of the Jacobian are the solution of $\lambda^2 - (Tr)\lambda + (Det) = 0$ using Trace (Tr) and Determinant (Det) for the matrix, the eigenvalues are

$$\lambda = [\,(Tr) \pm \sqrt{(Tr)^2 - 4(Det)}\,] \,/\, 2. \tag{11}$$

For the Hopf bifurcation, it must be satisfied that both eigenvalues of the Jacobian are imaginary. The condition is equivalent to $(Det) > 0$ at $(Tr) = 0$ (Guckenheimer & Holmes, 1983).

Theorem 2 *To make the Hopf bifurcation at a unique equilibrium point for one E-I pair, the conditions $K_{EI} > W_{ii}\pi x_i^*/2$ and $W_{ii} \geq \pi a$ are both necessary.*
Proof. First, the condition for $(Tr) = 0$ is discussed.
From the condition $(Tr) = -2 + W_{ii}G'(v_e) = 0$,

$$\frac{2}{W_{ii}} = G'(v_e) = \frac{2a}{\pi(a^2 + v_e^2)}, \tag{12}$$

is obtained, and the condition $(Tr) = 0$ is equivalent to $\pi(a^2 + v_e^2) = aW_{ii}$.
If $I_i > 0$, that is, $v_e > 0$, is assumed, using $v_e = \sqrt{a(W_{ii} - \pi a)/\pi}$,

$$x_i^* = G(v_e) = \frac{2}{\pi}\arctan\sqrt{\frac{W_{ii} - \pi a}{\pi a}} = const. \tag{13}$$

The term $v_e = I_i + W_{ii}x_i^* - K_{EI}y_i^*$ is put into the equation (12) , which is arranged on I_i.

$$-\pi I_i^2 - 2\pi(W_{ii}x_i^* - K_{EI}y_i^*)I_i - \pi(W_{ii}x_i^* - K_{EI}y_i^*)^2 + a(W_{ii} - \pi a) = 0. \tag{14}$$

To satisfy this condition, it is necessary that the quadratic equation (14) on I_i have real roots. By the discriminant $D = 4\pi a(W_{ii} - \pi a) \geq 0$, the condition $W_{ii} \geq \pi a$ is necessary.
Second, the condition for $(Det) > 0$ is discussed at $(Tr) = 0$.
By $(Tr) = -2 + W_{ii}G'(v_e) = 0$,

$$(Det) = 1 + G'(v_e)[\,-W_{ii} + K_{IE}K_{EI}G'(v_i)\,] \tag{15}$$
$$= -1 + 2K_{IE}K_{EI}G'(v_i)/W_{ii} > 0, \tag{16}$$

is equivalent to the condition $W_{ii} < 2K_{IE}K_{EI}G'(v_i)$.
The term $v_i = K_{IE}x_i^*$ is put into inequality (16), which is arranged on K_{IE},

$$W_{ii} < \frac{4aK_{EI}K_{IE}}{\pi(a^2 + (K_{IE}x_i^*)^2)}, \tag{17}$$

$$(W_{ii}\pi(x_i^*)^2)K_{IE}^2 - (4aK_{EI})K_{IE} + (W_{ii}\pi a^2) < 0. \tag{18}$$

To satisfy the inequality (18) on $K_{IE}(> 0)$, the discriminant

$$D = (4aK_{EI})^2 - 4(W_{ii}\pi(x_i^*)^2)(W_{ii}\pi a^2) = 16a^2[\,(K_{EI})^2 - (W_{ii}\pi x_i^*/2)^2\,] > 0, \tag{19}$$

is necessary.
Consequently, the condition $K_{EI} > W_{ii}\pi x_i^*/2$ is necessary. (QED).

Thus, under several assumptions, we have obtained two theoretical results: an asymptotical stable point near a memory for a sufficiently small K_{IE}, and the necessary condition of the Hopf bifurcation for one E-I pair, although they are references to the restricted behaviors near a memory.

4 DYNAMIC BEHAVIORS OF THE NETWORK

In this section, we show the main result of this paper using computer simulations. A limit cycle is generated for an input near a memory, while a chaotic orbit is generated for an input far from memories.

4.1 Simulation Results for One E-I pair

As a basic hehavior of the oscillatory neural network, the simulation result for only one E-I pair is shown below. The stability map for parameters I_i and K_{IE} is shown in Fig. 3. In order to satisfy the necessary conditions of the Hopf bifurcation of one E-I pair from Theorem 2, the parameters are set to $W_{ii} = 1.0, K_{EI} = 2.0, a = 0.1$. At the bifurcation point, whether an oscillation actually occurs depends on the input bias.
On the dotted line in Fig. 3, oscillation does not occur at some candidate points for the Hopf bifurcation because the candidate points become saddle points as $(Det) < 0$ at $(Tr) = 0$, and give birth to another stable point through bifurcation.
For a greater value of K_{IE}, the joining of two points with one unstable point causes an oscillation. This is not the same case as the Hopf bifurcation of Theorem 2 but rather a Birth-Death bifurcation. From the results of the simulation of one E-I pair, the maximum frequency is limited by the value of K_{IE}. When the value of K_{IE} is fixed ($K_{IE} = 1.0$), the frequency is increased by decreasing the value of I_i (Fig. 4). The initial states are set at $x_i(0) = I_i$ and $y_i(0) = 0$. Moreover, resonant phenomena can be found for a forced sinusoidal input.

4.2 Trajectories in the 3D-Pattern Space

The trajectories of the oscillatory neural network model are shown with an example of 3 E-I pairs. Each of the three memory patterns ξ^μ; $\xi^1 = (1,1,1)$, $\xi^2 = (1,-1,-1)$, $\xi^3 = (-1,-1,1)$, is set into the auto-correlation matrix. In order to satisfy the condition of Theorem 2, the parameters are set to $K_{EI} = 2.0$, $a = 0.1$, $(W_{ii} = 1,\ from\ N = M)$. The bifurcation parameter K_{IE} is variable.
The activation values of excitatory cells are explained by a point in the three-dimensional coordinates $\{x_i\}$. Each memory pattern is explained by the vertex of the cube; $-1 \leq x_i \leq 1\ (i = 1 \sim 3)$. An input bias $\{I_i\}$ for excitatory cells is given by a $(-1,1)$ uniform random pattern. The initial states are set at $x_i(0) = I_i$ and $y_i(0) = 0$.

When the value of K_{IE} is sufficiently small $(K_{IE} = 0.05)$, the asymptotical stability near a memory is shown in Fig. 5(a). The trajectories converge to an equilibrium point near each memory. Also, for the case of $K_{IE} = 0.3$, we show the trajectory after the Hopf bifurcation of the third pair $x_3 - y_3$ (Fig. 5 (b)). The results are references to the cases of Theorem 1 & 2, respectively.

At a larger value, $K_{IE} = 0.5$, we show examples of limit cycles as stable orbits (Fig. 6). For an input pattern near the memory pattern, the oscillation converges to a simple limit cycle. However, for a pattern far from the three memory patterns, the oscillation becomes a chaotic complex orbit. Though the limit cycle is almost independent of the initial states of the cells, the orbit begins to depend on the initial states as an input far from the memory patterns. Since another orbit exists for a slightly different input bias, there are many limit cycles in the pattern space.

It is shown that higher freqency waves exist for a large value of K_{IE} in Fig. 7. For the case of a small value of K_{IE} in Fig. 7 (a), there are simple low frequency waveforms of $x_i(t)$. However, for the case of a large value of K_{IE} in Fig. 7 (b), there are complex waveforms of $x_i(t)$ with high frequencies.

The distance between $x_i(t)$ and ξ^μ is evaluated by the overlap function $m^\mu(t)$ (Horn & Usher, 1991). The recognition result is given by the maximum peak value of $m^\mu(t)$ or the maximum fraction of the time τ; $m^\mu(\tau) >$ (*threshold value*).

$$m^\mu(t) = \frac{1}{N} \sum_{j=1}^{N} \xi_j^\mu x_j(t). \tag{20}$$

In the example in Fig. 8 (a), the orbit is a vibration between vertex ξ^2 and ξ^3. However, in Fig. 8 (b), the orbit is very complex and wanders among three memory pattern vertexes.

4.3 Synchronized Oscillations and Spectrum Analysis

For the previous example, the synchronization of cells and the chaotic character of the orbit are investigated for an input, whether near or far from the memory patterns.

The cross-correlograms (temporal cross-correlations) between the activations of excitatory cells are shown in Fig. 9. Three histograms are respectively related to $x_1 - x_2$, $x_2 - x_3$, $x_3 - x_1$, below. The auto-correlograms (temporal auto-correlations) for each activation; x_1, x_2, x_3, are shown in Fig. 10. For an input near a memory pattern, these activations are synchronous and make the limit cycle in Fig. 9 (a) and Fig. 10 (a). However, for an input far from the memory pattern, they are asynchronous in Fig. 9 (b) and Fig. 10 (b). Since the temporal correlations are vanishing, it is shown that future activities are unpredictable, which is one of the characteristics of chaos.

The chaotic characteristic of the orbit is also investigated by spectrum analysis for the waveform of the overlap function $m^1(t)$. In Fig. 11 (a), there are several sharp peaks for an input near a memory pattern. However, in Fig. 11 (b), a continuous spectrum with various frequencies exists in the waveform for an input far from the memory patterns. This continuous spectrum shows the chaotic character of the orbit. For another $m^2(t)$ and $m^3(t)$, the same results are obtained. The left-hand side of each spectrum is a DC element whose frequency is zero.

The occurrence of chaotic attractors near quasi-periodic orbits on the $m-$torus ($m \geq 3$) is suggested by Ruelle and Takens et al. (1971, 1978). If many oscillators (more than 3) each have an independent frequency and are weakly coupled, a chaotic orbit can occur.
In the proposed model, each E-I pair is an autonomous oscillator and is weakly coupled by the auto-correlation matrix $W_{ij} = O(1/N)$. Thus, for an input far from the memory patterns, it can be considered that the chaotic orbit bifurcates from the torus through the simple limit cycle by Ruelle-Takens's scenario.

The main result of this paper is that we obtained the following characteristic behavior.

- a limit cycle near a memory (memory retrieval with anbiguous fluctuation) for an input near a memory
- a chaotic orbit among memories (autonomous search) for an input far from memories

5 LEARNING OF CONTINUOUSLY TRANSFORMED PATTERN CYCLES

In this section, continuously transformed pattern cycles are embedded in the limit cycle of the oscillatory neural network by learning. It is considered that the physiological synchronized oscillations may be related to some kind of orderly process such as the cognitive psychological transformational operation of characters. The existence of such a transformation in the brain is discussed from the cognitive psychological point of view in the next section.
An example of continuous transformation for hand-written characters is explained. Then, an improvement of Sato's learning method is investigated in order to efficiently embed these continuously transformed patterns in the limit cycles of the oscillatory neural network.

5.1 Continuously Transformed Pattern

For various character patterns in a category, it is considered that the main factors of the distinctions are shift, and local distortion as a nonlinear transformation.
In a simulation of guessing age using a transformed face image, a continuous transformation of both shift and distortion has been used (Shaw & Wilson, 1976). The transformation is defined by the combination of shift transformation for the coordinates of the whole image by degree ϕ, and the following cardioid transformation,

$$r_{new} = r\,(\,1\,-\,k\sin\theta\,), \tag{21}$$

$$\theta_{new} = \theta. \tag{22}$$

Where, k is the parameter for the amount of distortion.
An example of a character pattern series by shift/distortion transformation is shown in Fig. 12. Although it is not clear whether something like this transformation exists in a pattern recognition process or perception in the brain, in this paper the transformation is assumed to be an example of the psychological continuous transformation of characters.
Of course, the transformation of direct images is used only for simplicity. More exactly, we must consider transformations not of the pattern images but of the internally represented patterns which are preprocessed by something like feature extractions.
Each transformed pattern cycle is used as a teacher signal for the excitatory cells. In the next subsection, we will discuss the learning method.

5.2 Learning Method Without Temporal Inverse Process

The character patterns used in this simulation are explained. For input patterns, Japanese hiragana characters of three categories: *a*, *i*, *u*, are used from the ETL9 hand-written Japanese character database. Each binary pattern (64x63 pixels) from the ETL9 database is converted to a mesh density pattern (16x16 = 256 gray levels) after normalization of size. These patterns are shown in Fig. 13.
Thus, the oscillatory neural network has 256 E-I pairs. Continuously transformed density patterns will be cyclically reconstructed by the network after learning.

For a recurrent neural network with feedback loops, a learning algorithm has been proposed (Sato, 1990; Pearlmutter, 1989). If the learning algorithm is directly applied to this network, the values of the bifurcation parameter $K_{IE}^{(i)}$ are updated through the temporal inverse process. Changing the bifurcation parameters gives rise to a loss of structural stability, and an increase in computation time to handle the iterations.

When the value of $K_{IE}^{(i)}$ is fixed as a constant value K_{IE}, the learning method becomes efficient without the temporal inverse process [1]. However, we must note that the fixing of K_{IE} restricts the ability of the network.

The improved learning method is processed as follows.
First, three memory patterns are defined by each averaged sample pattern (each element is converted to a binary element, ± 1) in each category. The values of connection weight coefficients W_{ij} are initialized by the auto-correlation matrix using these memory patterns, and both K_{IE} and $K_{EI}^{(i)}$ are initialized to satisfy the oscillatory condition of Theorem 2 ($K_{IE} = K_{EI}^{(i)} = 2.0$, $a = 0.1$).
By the effect of the auto-correlation matrix, since the trajectory for an input near a memorized character pattern may not converge to a stable point or chaotic attractor, but to a limit cycle near the memory point, it is possible that the learning of the transformational cycles from the limit cycle near the memory point does not need to alter the structural stability.

Second, the learning of a continuously transformed pattern is iterated using the teacher signals for excitatory cells. The teacher pattern cycles are created by the shift/distortion transformation. In the learning of these connection coefficients, one cycle of the teacher pattern series is divided into a number of sections.

On each section, the teacher pattern is given continuously by a linear (or spline) complement between two transformed binary patterns. In the learning for a section, each excitatory cell is clamped by the teacher signal, and the connection weight coefficients are updated by the steepest descent direction method for the error function E.
In order to equalize the activation value $x_i(t)$ to the teacher signal $Q_i(t)$, the forced input $J_i(t)$ is defined as follows ($J_i(t) \to 0$; $x_i(t) \to Q_i(t)$).

$$E = \frac{1}{2} \int_T^{T+TB} \sum_{i=1}^{N} [\, J_i(t) \,]^2 dt, \tag{23}$$

$$J_i(t) = \dot{Q}_i(t) + Q_i(t) - G(v_e^{(i)}(t)), \tag{24}$$

$$v_e^{(i)}(t) = \sum_{i=1}^{N} W_{ij} Q_i(t) - K_{EI}^{(i)} y_i(t) + I_i, \tag{25}$$

$$\Delta W_{ij} = -\eta \frac{\partial E}{\partial W_{ij}}. \tag{26}$$

Where, $[T,\ T+TB]$ is the time required for one section of the learning. η is a learning coefficient.

According to the calculations of Sato's learning algorithm (Sato, 1990), the learning equations are obtained as follows (see Appendix). Where $P_i(t)$ is a Lagrange multiplier.

$$P_i(t) = -J_i(t), \tag{27}$$

$$\Delta W_{ij} = -\eta \int_T^{T+TB} [\, P_i(t) G'(v_e^{(i)}(t)) x_j(t) \,] dt, \tag{28}$$

$$\Delta K_{EI}^{(i)} = \eta \int_T^{T+TB} [\, P_i(t) G'(v_e^{(i)}(t)) y_i(t) \,] dt. \tag{29}$$

The learning for a pattern cycle is integrated with updatings of some sections, and these processes are also iterated for all learning sample patterns until reaching a small error value.

[1] personal communication with Dr. Masa-aki Sato of ATR Human Information Processing Research Labs.

5.3 Simulation Results

As the memory patterns, three handwritten hiragana characters; a, i, u are shown in Fig. 13. Each pixel is displayed using 256 gray levels.

In Figs. 14 and 15, the output pattern series before learning are shown from upper-left to lower-right. In this case, the connections W_{ij} are defined by the auto-correlation matrix. For a character input, the pattern series in Fig. 14 is periodically changing near the memory pattern, though many of them are not read as the characters.

For a random input, the pattern series in Fig. 15 is chaotic while autonomously searching among the memory patterns and the white-black inverse patterns. We can see the transition; inverse u (27-29th) $\rightarrow$ u (35-38th) $\rightarrow$ inverse i (46-49th) $\rightarrow$ i (55-58th) $\rightarrow$ a (60-62th) $\rightarrow$ inverse i (65-67th) $\rightarrow$ inverse a (69-70th).

However, the output pattern series after learning each pattern with 1000 iterations is shown in Fig. 16. It is shown that the teacher pattern cycle in Fig. 12 is embedded in the limit cycle. For another two categories or characters, the same results are obtained.

In contrast, for a random input after learning, the output series shows a chaotic search similar to the series before learning. The integrated error in each cycle is converged before 1000 iterations.

6 SUMMARY AND DISCUSSION

We have investigated experimentally the dynamic behaviors of an oscillatory neural network with E-I pairs, and also analyzed theoretically a few restricted behaviors near a memory in the state space. Computer simulations of a three-dimensional case have shown an interesting characteristic: the autonomous generation of a limit cycle near a memory for an input near a memory, and of a chaotic orbit among memories for an input far from memories. The purposes of information processing include memory retrieval with ambiguous fluctuation and autonomous search among memories, respectively.

Thus, the function of the network is not fixed, but drastically changeable according to the relation between an input and memories (distance or correlation, etc.). In order to realize dynamic information processing for a spatiotemporal pattern in the brain, this changeability may be one of the advantages of such a nonlinear, non-equilibrium system.

As an example of dynamic information processing, it has been shown that continuously transformed pattern cycles for three Japanease characters are embedded in the limit cycles of the oscillatory neural network by the improved Sato's learning method without the temporal inverse process. In this higher-dimensional case, we show the characteristic orbit (limit cycle or chaos which changes according to the input). However, there are something about dynamic information processing in the brain that are unclear.

In the following, we will discuss the continuous transformation of perceptions from the cognitive psychological point of view.

It is considered that a human will use conceptions not only of the invariant amounts, but also of the continuity of the transformation, to perceive various objects as the same pattern or concept (Piaget, 1970). For example, mental rotation is a well known process of inner-pattern transformation in the human brain.

Of course, the conventional process of using features is not denied; both mechanisms may be selectively or cooperatively processed according to the situation. Takano (1989) has reported that if there are obvious distinctions of the elements or the relations between the line segments of figures, mental rotation is not performed and the similarity is decided by the distinctions of the features.

On Piaget's genetic epistemology (Piaget, 1970; Sigel & Cocking, 1977), the construction of the

concept of conservation for material amounts is shown, and these amounts can be operated by a child after the age of $7 \sim 8$. In Piaget's theory, which is based on considerable experience in the field of psychology, the following are expressed.

1. At this stage (age $7 \sim 8$), a mental image is changed from a static image to a dynamic image for motion or transformation using operations.

2. Since invertible and elastic sensual-motional operations can be used, the expected image becomes more concrete (e.g. for a relation between forms and amounts to a transformation from "sausage" to "ball").

3. After this stage, the concepts of conservation, quantity, categorization and sequential formulation are acquired by using the inner continuous operations, and functional images are constructed only by perception.

In relation to Piaget's theory, Tanaka (1969) has reported that the conceptual and operational cognition stage starts from the age of $7 \sim 10$ by experimental tests for the similarity of figures. It is shown that a child (age $3 \sim 4$) discriminates the similarity from only a few local features by ignoring the global form. But an older child (age $8 \sim 9$) does so using operations. The results are not only unexplainable by conventional recognition methods based on feature extractions, but also show the existence of inner-transformations or operations.

However, Yokose (1986) has reported interesting results for recognition of the alphabet and Japanese katakana characters by testing subjects ranging from elementary school students to college students. The response time is proportional to the degree of distinction and the recognition rate is proportional to age. It is believed that this is because an inner transformation takes considerable time for a greater distinction and some students are not yet able to do this. These results are relevant to Piaget's theory.

Another reason for the existence of the inner transformations is supported by the transformational structure theory of pattern cognition (Ito, 1975; Imai, 1986). In this theory, the judgements of pattern "similarity" and "goodness" are determined after selecting and operating a cognitive pattern transformation. They pointed out that the recognition of transformation structures is an essential factor in determining perceptual classification or categorization. The selection in the transformational structures corresponds to whether mental rotation has to be performed in Takano's theory.

Consequently, it is believed that an inner continuous transformation is used in the pattern recognition process. However, it is not yet clear what kind of transformation is processed in the brain. To clarify this, we must investigate the inherent transformational operations in the recognition of various character patterns. Also, the ability of the oscillatory neural network to handle spatiotemporal pattern processing will be investigated in further study.

References

[1] Amari, S. (1971). Characteristics of Randomly Connected Threshold-Element Networks and Network Systems. *Proc. of the IEEE*, 59(1), 35-47.

[2] Amari, S. (1972a). Characteristics of Random Nets of Analog Neuron-Like Elements. *IEEE Trans on Sys, Man and Cybern*, SMC-2(5), 643-657.

[3] Amari, S. (1972b). Learning and Patterns and Pattern Seqences by Self-Organizing Nets of Threshold Elements. *IEEE Trans. on Computer*, c-12(11), 1197-1206.

[4] Baird, B. (1986). Nonlinear dynamics of pattern formation and pattern recognition in the rabbit olfactory bulb. *Physica*, 22D, 150–175.

[5] Doya, K. (1992). Bifurcation in the learning of Recurrent Neural Networks. *Proc. of IEEE Int. Symp. on Circuits and Systems*, 2777-2780.

[6] Eckhorn, R., Bauer, R., Jordan, W., Brosch, M., Kruse, W., Munk, M. & Reitboeck, H. J. (1988). Coherent oscillations: A mechanism of feature linking in the visual cortex? - multiple electrode and correlation analysis in the cat -. *Biological Cybernetics*, 60, 121–130.

[7] Eckhorn, R., Reitboeck, H. J., Arndt, M. & Dicke, P. (1990). Feature Linking via Synchronization Among Distributed Assemblies: Simulations of Results from Cat Visual Cortex. *Neural Computation*, 2(3), 293–307.

[8] Gray, C. M. & Singer, W. (1989). Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proc. Natl. Acad. Sci. USA*, 86, 1689–1702.

[9] Gray, C. M., Kőnig, P., Engel, A. K. & Singer, W. (1990). Synchronization of Oscillatory Responses in Visual Cortex: A Plausible Mechanism for Scene Segmentation. *In H. Haken and M. Stadler (Eds), Synergetics of Cognition*, volume 86, (pp. 82–98). Springer-Verlag.

[10] Grossberg, S. & Somers, D. (1991). Synchronized oscillations during cooperative feature linking in a cortical model of visual perception. *Neural Networks*, 4(4), 453–466.

[11] Guckenheimer, J. & Holmes, P. (1983). *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields.* Springer-Verlag.

[12] Horn, D. & Usher, M. (1991). Parallel activation of memories in an oscillatory neural network. *Neural Computation*, 3(1), 31–43.

[13] Imai, S. (1986). The Transformational Structure Theory of Pattern Cognition. *Japanese Psychological Monographs*, (17). The Japanese Psychological Association.

[14] Ito, S. (1975). Similarity judgements of patterns depending upon perceived inter-pattern transformation structure. *The Japanese Journal of Psychology*, 46(1), 10–18.

[15] Kőnig, P. & Schillen, T. B. (1991). Stimulus-dependent assembly formation of oscillatory responses: I. synchronization and II. desynchronization. *Neural Computation*, 3(2), 155–177.

[16] Kuramoto, Y. (1991). Collective synchronization of pulse-coupled oscillators and excitable units. *Physica*, D50, 15–30.

[17] Lie, Z. & Hopfield, J. J. (1989). Modeling the olfactory bulb and its neural oscillatory processings. *Biological Cybernetics*, 61, 379–392.

[18] Lummer, E. & Huberman, B. A. (1992). Binding Hierarchies: A Basis for Dynamic Perceptual Grouping. *Neural Computation*, 4(3), 341–355.

[19] Matsuoka, K. (1987). Mechanisms of frequency and pattern control in the neural rhythm generators. *Biological Cybernetics*, 56, 345–353.

[20] Mori, Y., Davis, P. & Nara, S. (1989). Pattern retrieval in an asymmetric neural network with embedded limit cycles. *J. Phys. A: Math. Gen.*, 22, L525–L532.

[21] Newhouse, S., Ruelle, D. & Takens, F. (1978). Occurrence of Strange Axion A Attractors Near Quasi-Periodic Flows on $T^m, m \geq 3$. *Communications in Mathematical Physics*, 64, 35–40.

[22] Niebur, E., Kammen, D. M. & Koch, C. (1991). Phase-Locking in 1-D and 2-D Networks of Oscillating Neurons. *In H. G. Schuster (Eds), Nonlinear Dynamics and Neuronal Networks - Proceedings of the 63rd W. E. Heraeus Seminar Friedrichsdorf 1990* -, (pp. 173–203). VCH.

[23] Pearlmutter, B. A. (1989). Learning state space trajectories in recurrent neural network. *Neural Computation*, 1(2), 263–269.

[24] Piaget, J. (1970). Piaget's theory. *In P. Mussen (Eds), Carmichael's manual of child psychology*, volume 1, (pp. 703–732). New York: Wiley.

[25] Reitboeck, H. J., R. Eckhorn, R., Arndt, M. & Dicke, P. (1990). A Model for Feature Linking via Correlated Neural Activity. *In H. Haken and M. Stadler (Eds), Synergetics of Cognition*, volume 86, (pp. 112-125). Springer-Verlag.

[26] Renals, S. & Rohwer, R. (1990). A study of network dynamics. *Journal of Statistical Physics*, 58, 825–848.

[27] Ruelle, D. & Takens, F. (1971). On the Nature of Turbulence. *Communications in Mathematical Physics*, 20, 167–192.

[28] Sakarda, C. A. & Freeman, W. J. (1987). How brains make chaos in order to make sense of the world. *Behavioral and Brain Sciences*, 10, 161–195.

[29] Sato, M. (1990). A learning algorithm to teach spatiotemporal patterns to recurrent neural networks. *Biological Cybernetics*, 62, 259–263.

[30] Sato, M., Murakami, Y. & Joe K. (1990a). Chaotic dynamics by recurrent neural networks. *Proceedings of the International Conference on Fuzzy Logic and Neural Networks*, 601–604.

[31] Sato, M., Joe, K. & Hirahara T. (1990b). APOLONN brings us to the real world : Learning nonlinear dynamics and fluctuations in nature. *Proc. of IJCNN '90*, I581-I587.

[32] Schuster, H. G. & Wagner, P. (1990). A model for neuronal oscillations in the visual cortex - 1. mean-field theory and derivation of the phase equations -. *Biological Cybernetics*, 64, 77–82.

[33] Shaw, R. & Wilson, B. E. (1976). Abstract conceptual knowledge: How we know what we know. *In D. Klahr (Eds), Cognition and instruction*, (pp. 197–221). Lawerense Erlbraum.

[34] Siegel, I. E. & Cocking, R. R. (1977). *Cognitive Development from Childhod to Adolescence: A Constructivist Perspective.* Holt, Rinehart and Winston.

[35] Sompolinsky, H., Golomb, D & Kleinfeld, D. (1991). Phase Coherence and Computation in a Neural Network of Coupled Oscillators. *In H. G. Schuster (Eds), Nonlinear Dynamics and Neuronal Networks - Proceedings of the 63rd W. E. Heraeus Seminar Friedrichsdorf 1990* -, (pp. 113–130). VCH.

[36] Takano, Y. (1989). Perception of rotated forms: A theory of information types. *Cognitive Psychology*, 21, 1–59.

[37] Tanaka, T. (1969). Analysis of perceptual and conceptual stages. *The Japanese Journal of Psychology*, 40(3), 146–150.

[38] Tsuda, I., Koerner, E. & Shimizu, H. (1987). Memory dynamics in asynchronous neural networks. *Progress of Theoretical Physics*, 78(1), 51–71.

[39] Wang, D., Buhmann, J. & von der Malsburg, C. (1990). Pattern segmentation in associative memory. *Neural Computation*, 2(1), 94–106.

[40] Wang, L., Pichler, E. E. & Ross, J. (1990). Oscillations and chaos in neural networks: An exactly solvable model. *Proc. Natl. Acad. Sci. USA*, 871, 9464–9471.

[41] Wilson, H. R. & Cowan, J. D. (1973). A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik*, 13, 55–80.

[42] Yao, Y. & Freeman, W. J. (1990). Model of biological pattern recognition with spatially chaotic dynamics. *Neural Networks*, 3(2), 153–170.

[43] Yokose, Z. (1986). *Psychology of Figure*, in Japanese, chapter 4, (pp. 88–144). Nagoya University Publisher.

A Appendix: Derivation of the Learning Equations

First, both N variables $x_i(t)$ and $y_i(t)$ are denoted by $2N$ variables $z_k(t)$, and other variables are also renewed using the same names.

$$\dot{z}_k(t) = -z_k(t) + G(u_k(t)), \tag{30}$$

$$u_k(t) = \sum_{j=1}^{2N} W_{kj} z_j(t) \quad (k \notin V). \tag{31}$$

$$z_k(t) = Q_k(t), \tag{32}$$

$$J_k(t) = \dot{Q}_k(t) + Q_k(t) - G(u_k(t)), \tag{33}$$

$$u_k(t) = \sum_{j=1}^{2N} W_{kj} z_j(t) + I_k \quad (k \in V). \tag{34}$$

Where V denotes the set of visible units, which are clamped by the teacher signal $Q_k(t)$.

Using the Lagrange multiplier $P_k(t)$, the error $E(W)$ is defined.

$$E(W) = \int_{T1}^{T2} dt \left[\frac{1}{2} \sum_{k \in V} J_k^2 - \sum_{k \in V} P_k \{\dot{Q}_k + Q_k - G(u_k) - J_k\} - \sum_{k \notin V} P_k \{\dot{z}_k + z_k - G(u_k)\} \right]. \tag{35}$$

The variation $\delta E(W)$ by δW is derived as follows.

$$\begin{aligned} \delta E(W) &= \textstyle\sum_{k \in V} \frac{\partial E}{\partial J_k} \delta J_k + \sum_{k \notin V} \frac{\partial E}{\partial z_k} \delta z_k + \sum_k \frac{\partial E}{\partial P_k} \delta P_k + \sum_{kj} \frac{\partial E}{\partial W_{kj}} \delta W_{kj} \\ &= \textstyle\int_{T1}^{T2} dt \, [\sum_{k \in V} (J_k + P_k) \delta J_k \\ &\quad - \sum_{k \notin V} \{P_k (\delta \dot{z}_k + \delta z_k) - \delta z_k \sum_j P_j G'(u_j) W_{jk}\} \\ &\quad + \sum_{k \in V} \delta P_k \{\dot{Q}_k + Q_k - G(u_k) - J_k\} + \sum_{k \notin V} \delta P_k \{\dot{z}_k + z_k - G(u_k)\} \\ &\quad + \sum_{kj} P_k G'(u_k) z_j \delta W_{kj} \,]. \end{aligned} \tag{36}$$

The variation of each variable by δW is defined as follows.

$$\delta J_k(t \mid W) = J_k(t \mid W + \delta W) - J_k(t \mid W) \quad (k \in V), \tag{37}$$

$$\delta z_k(t \mid W) = z_k(t \mid W + \delta W) - z_k(t \mid W) \quad (k \notin V), \tag{38}$$

$$\delta P_k(t \mid W) = P_k(t \mid W + \delta W) - P_k(t \mid W). \tag{39}$$

The third term of equation (36) is cancelled out by equations (30) and (33). Assume that the Lagrange multiplier $P_k(t)$ satisfies the following equations. By the following definition, the first term of equation (36) is also cancelled out.

$$\begin{aligned} P_k &= -J_k && (k \in V), \\ \dot{P}_k &= P_k - \textstyle\sum_j P_j G'(u_j) W_{jk} && (k \notin V). \end{aligned} \tag{40}$$

The second term of equation (36) is also cancelled out using the initial and boundary conditions of the integration.

$$\begin{aligned} [\textit{the second term}] &= -\int_{T1}^{T2} dt \sum_{k \notin V} (P_k \delta \dot{z}_k + \dot{P}_k \delta z_k) \\ &= -\sum_{k \notin V} [P_k \delta z_k]_{T1}^{T2} \\ &= \sum_{k \notin V} \{ P_k(T1 \mid W) \delta z_k(T1 \mid W) - P_k(T2 \mid W) \delta z_k(T2 \mid W) \} = 0. \end{aligned} \tag{41}$$

Consequently, only the fourth term remains. The learning equations are obtained.

$$\delta E(W) = \int_{T1}^{T2} dt \sum_{kj} P_k G'(u_k) z_j \delta W_{kj}, \tag{42}$$

$$\Delta W_{kj} = -\eta \frac{\partial E}{\partial W_{kj}} = -\eta \int_{T1}^{T2} dt [\ P_k G'(u_k) z_j\ ]. \tag{43}$$

Next, return the variables $z_k(t)$ to variables $x_i(t)$ and $y_i(t)$, and rewrite the learning equations. In order to reduce the temporal inverse processes of $\dot{P}_i(t)$, the value of $K_{IE}^{(i)}$ must be fixed.
If the value of $K_{IE}^{(i)}$ is fixed, the Lagrange multiplier is unnecessary for hidden unit $y_i(t)$ from equation (43). However, this means that the ability of the oscillatory neural network is restricted.

Thus, the learning equations of connection weight coefficients W_{ij} and $K_{EI}^{(i)}$, which connect into visible units (excitatory units), are obtained as follows. This process is efficient without the temporal inverse process of the Lagrange multiplier.

$$P_i(t) = -J_i(t), \tag{44}$$

$$\Delta W_{ij} = -\eta \int_{T1}^{T2} [\ P_i(t) G'(v_e^{(i)}(t)) x_j(t)\ ] dt, \tag{45}$$

$$\Delta K_{EI}^{(i)} = \eta \int_{T1}^{T2} [\ P_i(t) G'(v_e^{(i)}(t)) y_i(t)\ ] dt. \tag{46}$$

B FIGURES AND CAPTIONS

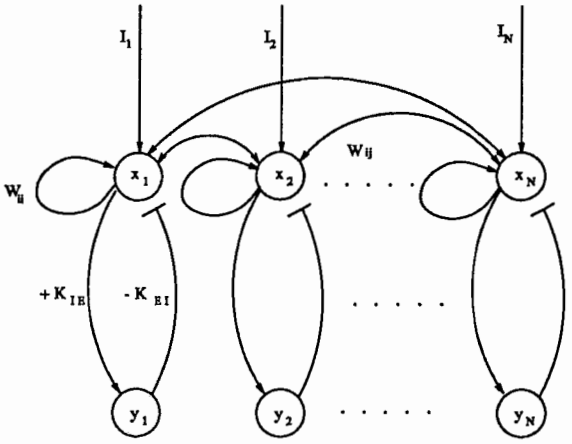


Figure 1: Oscillatory neural network. x_i and y_i are excitatory and inhibitory cells, respectively. $\dashv$ and $\rightarrow$ denote inhibitory connections and excitatory connections, respectively.

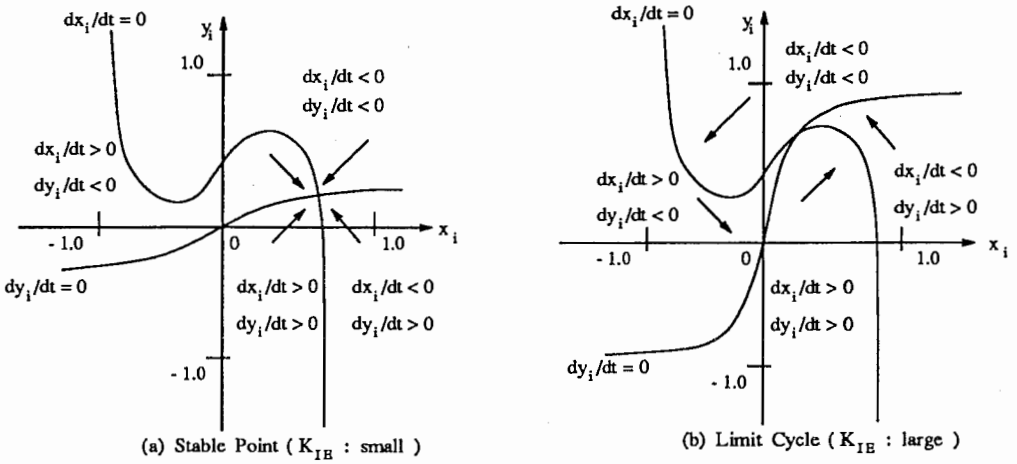


Figure 2: Vector field around the equilibrium point in one E-I pair. (a) When the value of K_{IE} is sufficiently small, there is only one asymptotical stable point (for a large input bias). (b) However, an oscillation occurs by increasing K_{IE}.

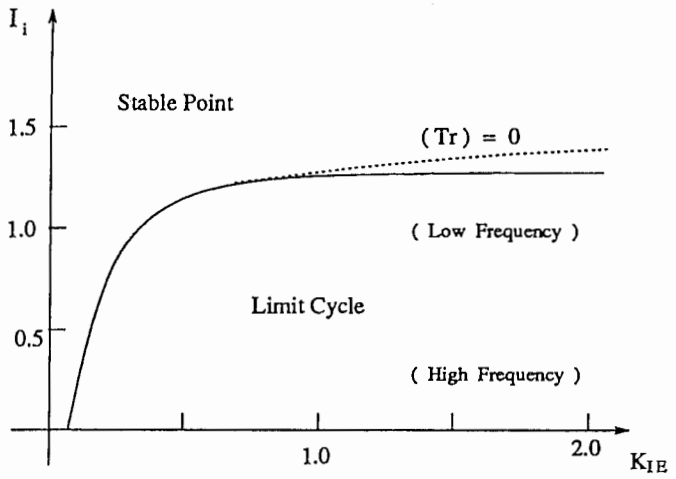
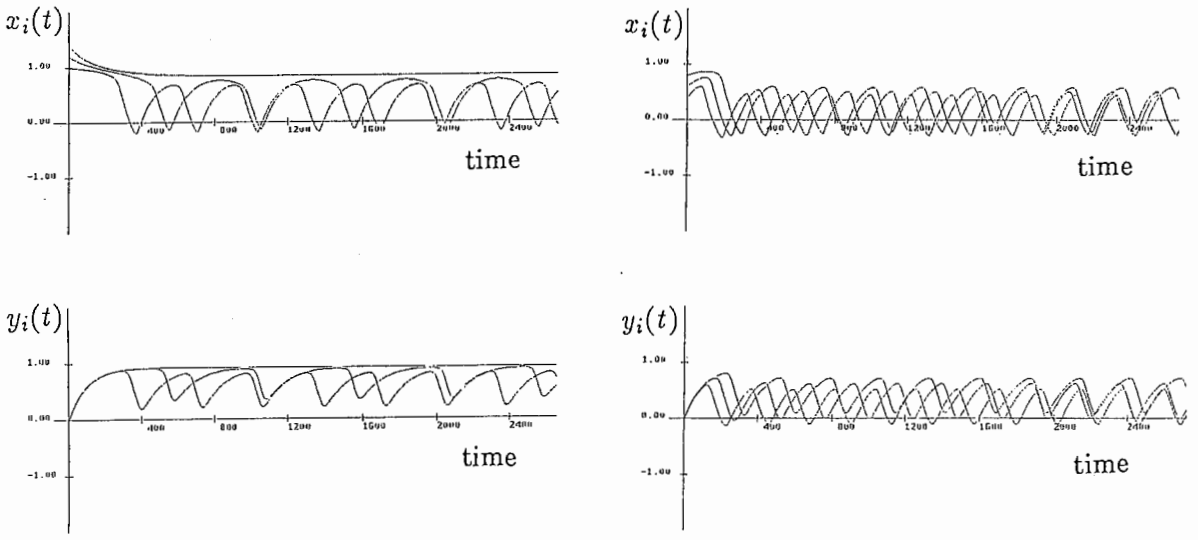
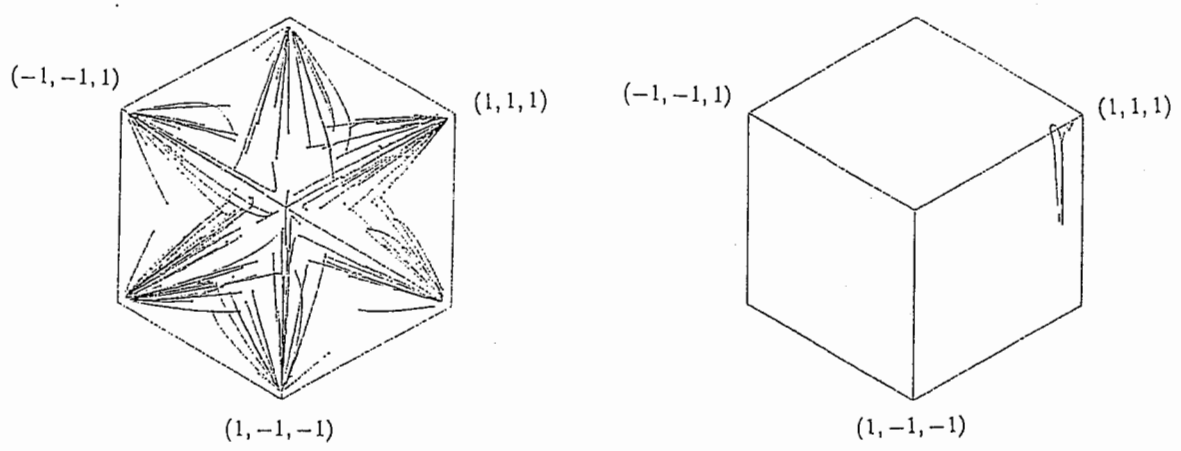


Figure 3: Stability map for parameters I_i and K_{IE}. The solid line denotes the critical line between the stable (or saddle) point and limit cycle. The dotted line denotes the no-oscillation case at the candidate of the Hopf bifurcation.



(a) $I_i = 1.4 \sim 1.0$ (b) $I_i = 0.8 \sim 0.4$

Figure 4: Waveforms of $x_i(t)$ and $y_i(t)$ for the input biases I_i of one E-I pair. The frequency is increased by decreasing the value of input bias I_i.



(*a*) $K_{IE} = 0.05$ (*b*) $K_{IE} = 0.3$

Figure 5: Examples of asymptotical stability and Hopf bifurcation. (a) Convergence of trajectories to the asymptotical stable points near memories. (b) Limit cycle oscillation after the Hopf bifurcation of the third pair $x_3 - y_3$.

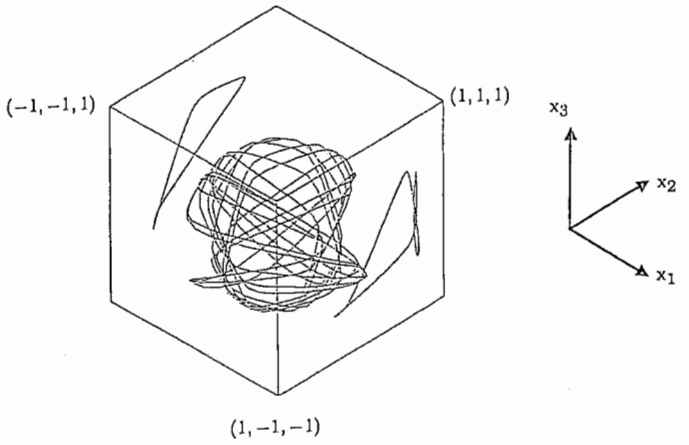
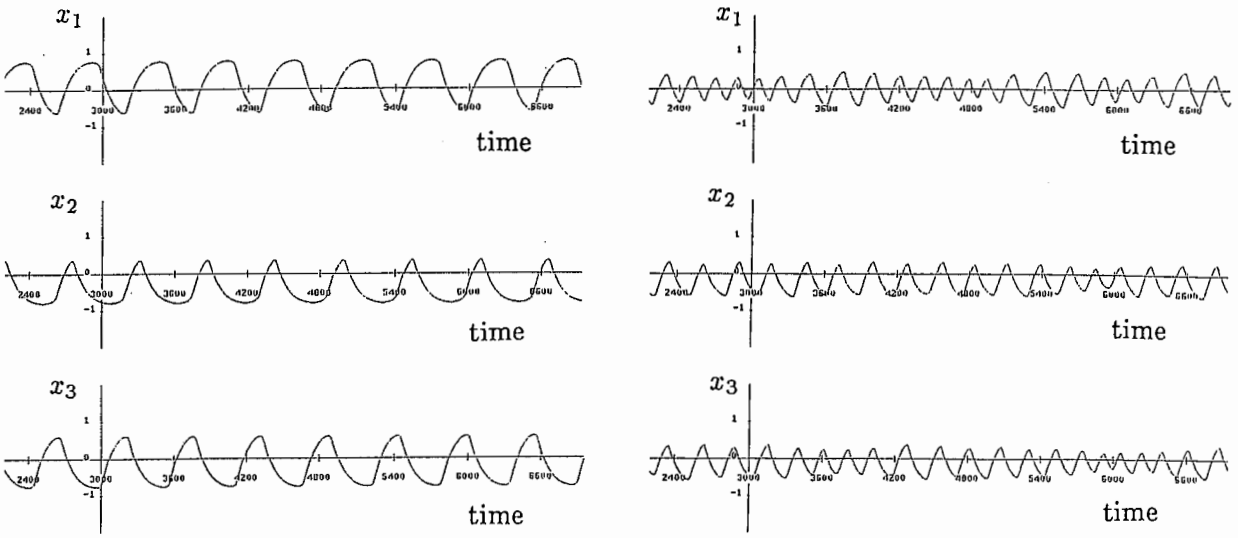
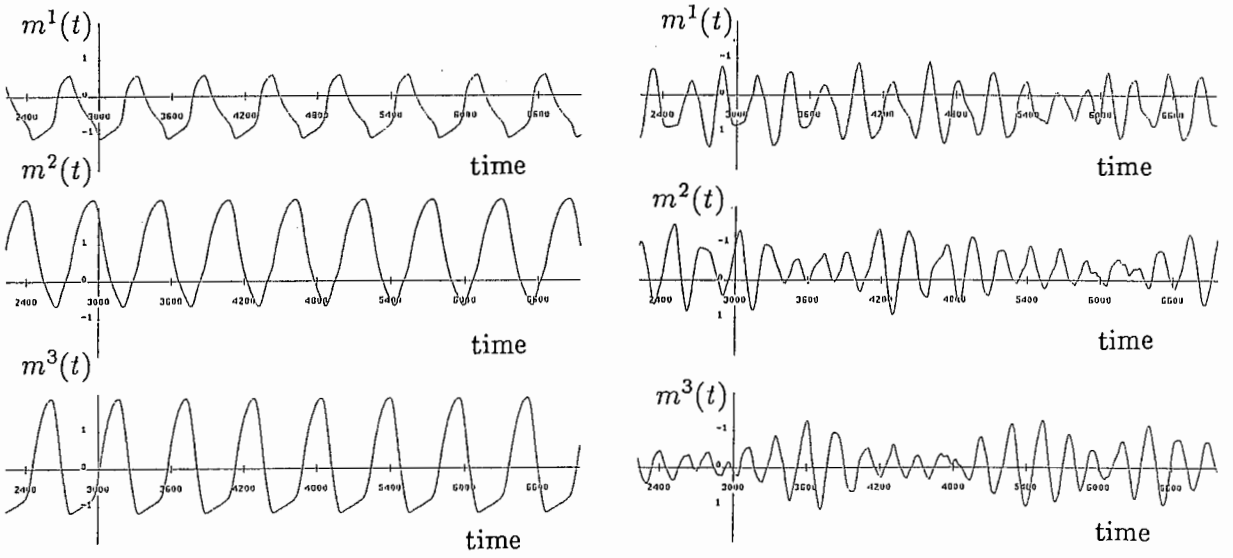


Figure 6: Two limit cycles and a chaotic orbit in the pattern space ($K_{IE} = 0.5$). In this figure, three orbits are superimposed on the same space. However, the dynamics in phase space is drastically changed (e.g. limit cycle $\leftrightarrow$ chaos) by the input pattern.



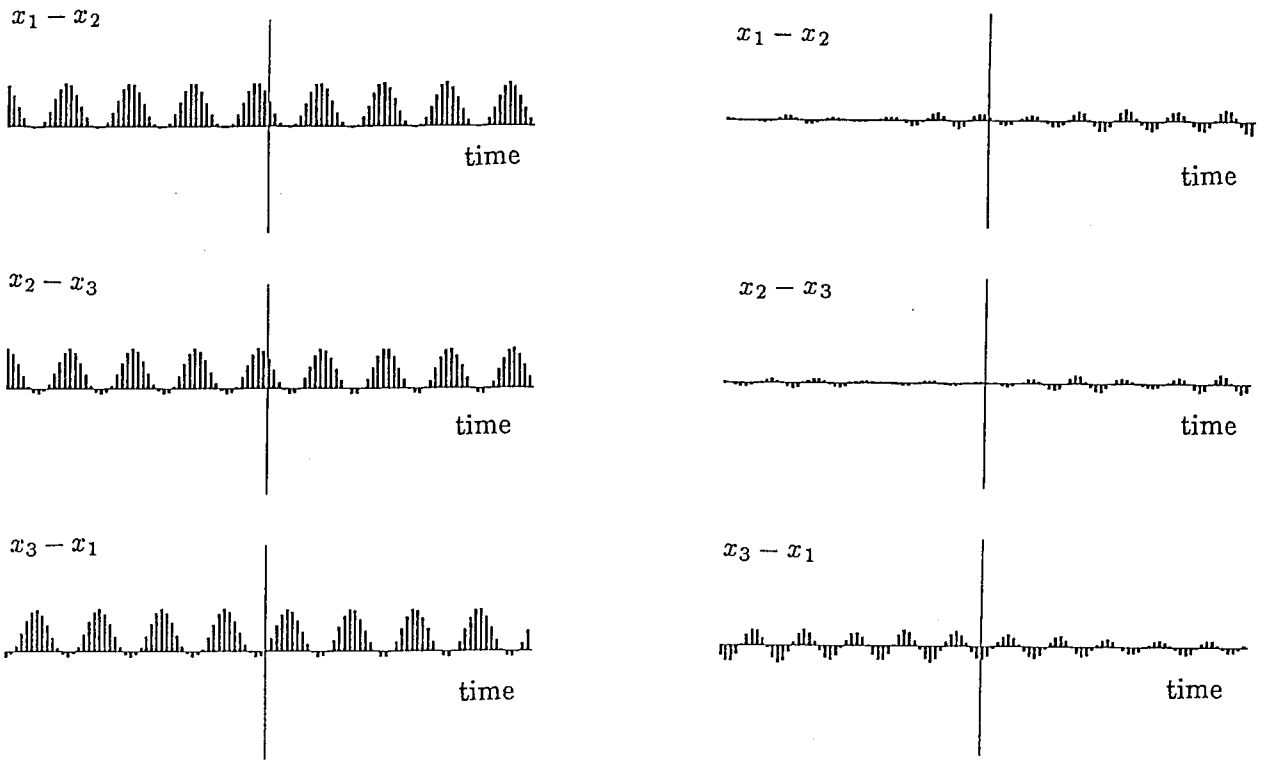
(*a*) *small* : $K_{IE} = 0.2$ (*b*) *large* : $K_{IE} = 2.0$

Figure 7: Time course of the activation $x_i(t)$.



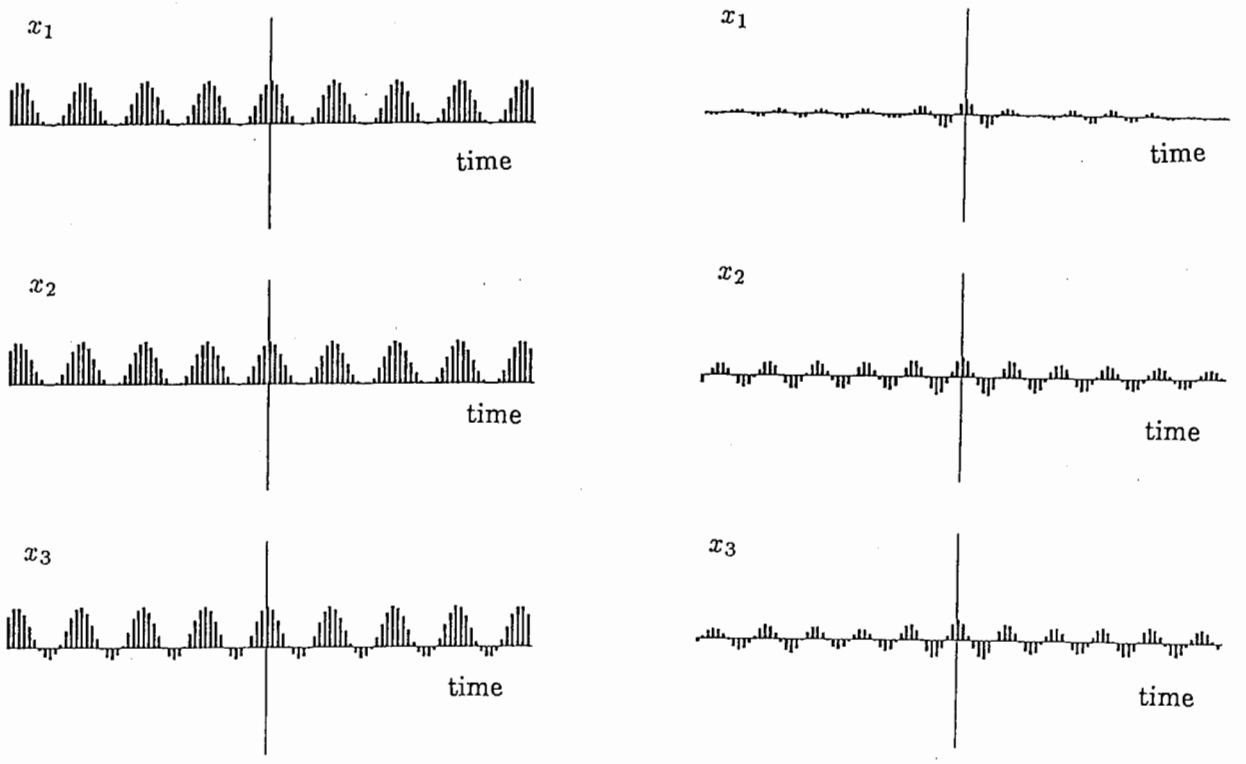
(*a*) *small* : $K_{IE} = 0.2$ (*b*) *large* : $K_{IE} = 2.0$

Figure 8: Time course of the overlap function $m^{\mu}(t)$.



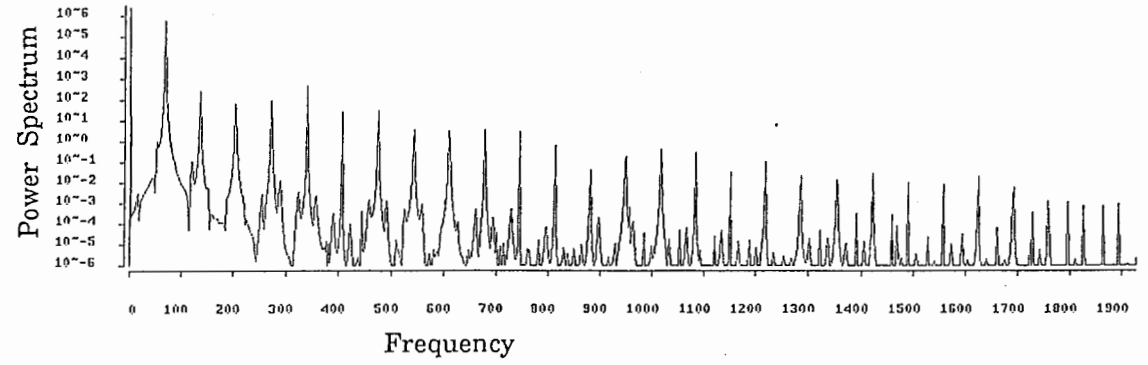
(*a*) *input near a memory pattern* (*b*) *input far from the memory patterns*

Figure 9: Cross-correlograms between the activations $x_i(t) - x_j(t)$. Each histogram denotes the correlation at a given time.

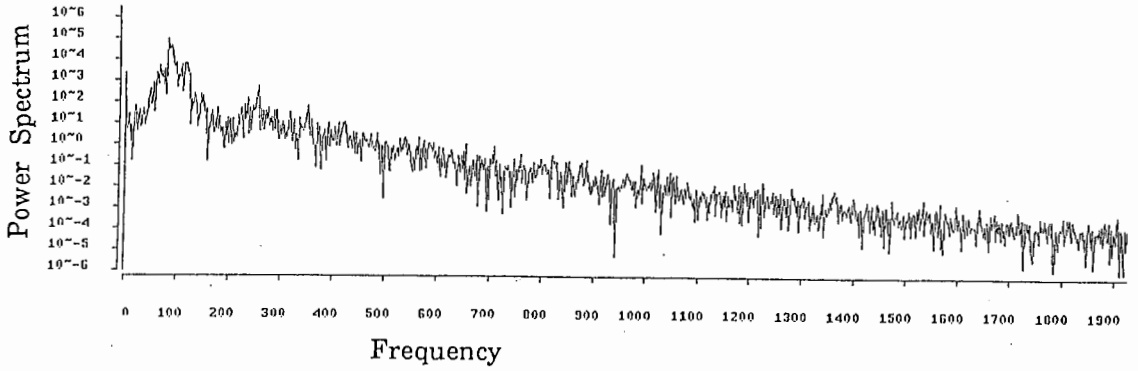


(*a*) *input near a memory pattern* (*b*) *input far from the memory patterns*

Figure 10: Auto-correlograms for the activation $x_i(t)$.



(*a*) *input near a memory pattern*



(*b*) *input far from the memory patterns*

Figure 11: Power Spectrum for the waveform of $m^1(t)$. (a) Several sharp peaks of a harmonic limit cycle. (b) Continuous spectrum with various frequencies, as chaotic character.

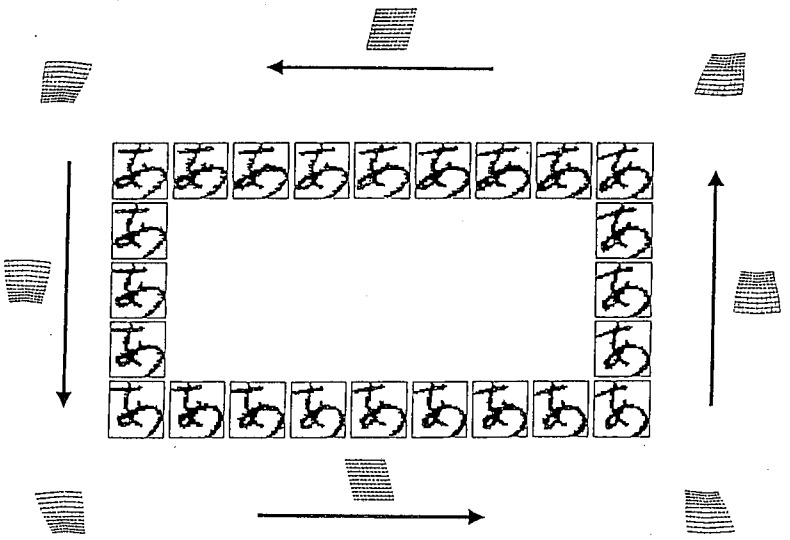


Figure 12: Example of a continuously transformed pattern cycle.

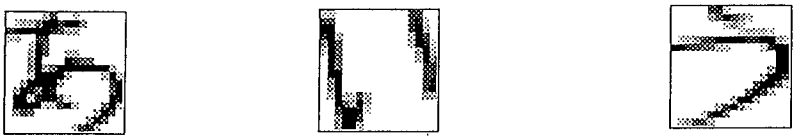


Figure 13: Three hand-written hiragana character patterns: *a*, *i*, *u*. Each pixel is displayed using 256 gray levels (from the value -1: white to +1: black).

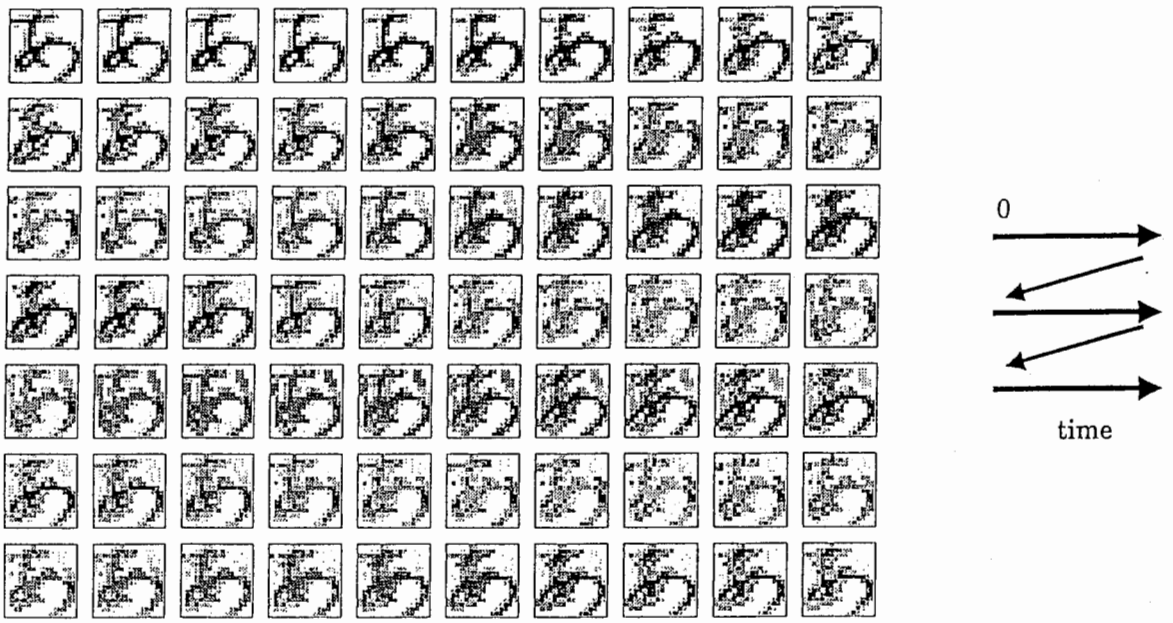


Figure 14: Output series for a character input before learning. These change periodically near the memory pattern, though many of them are not read as the character.

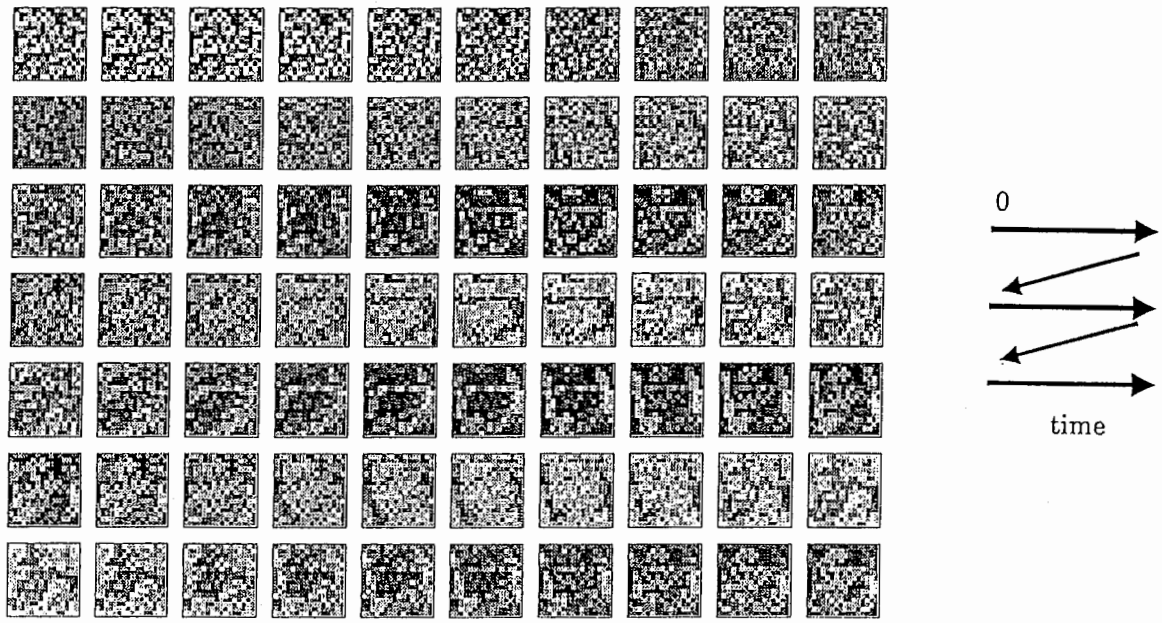


Figure 15: Output series for a random input before learning. These show autonomously chaotic searching among the memory patterns and the white-black inverse patterns.

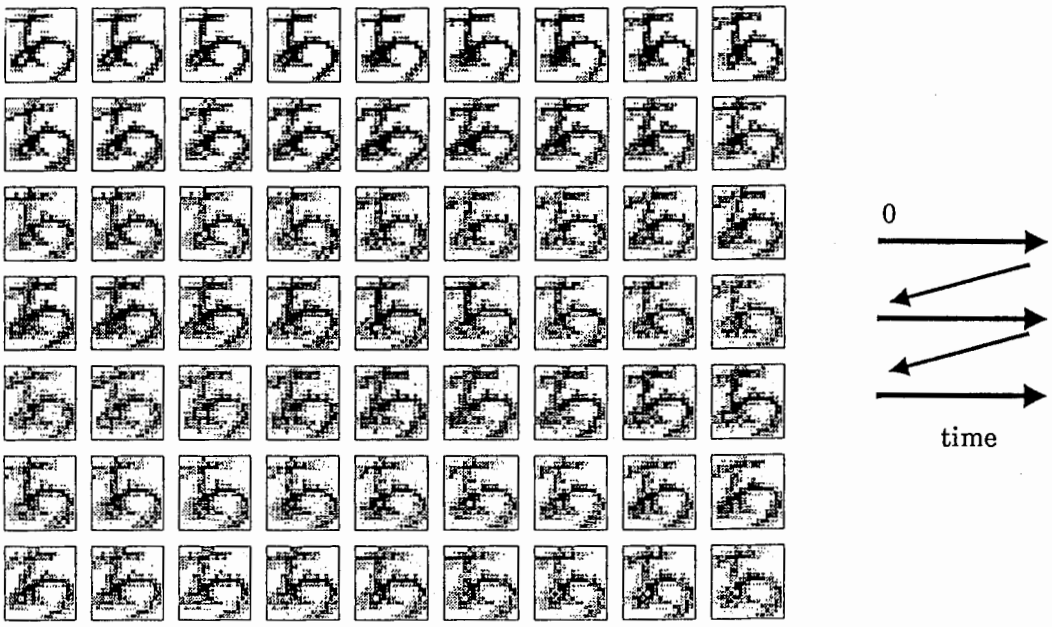


Figure 16: Output series for a character input after learning ($\eta = 0.0001$). These show the transformed pattern cycle of the teacher signals.