Journal of Statistical Physics, Vol. 6, No. 2/3, 1972

Time Partition Function Analysis of a Neural Network

Makoto Oonuki[1,2]

Received December 16, 1971

The time characteristics of a linear network in the brain are obtained by the method of the "time partition function," which is analogous to a grand partition function or a distribution function in statistical mechanics. The analogy between the average density in a many-particle system and the reciprocal of the frequency in a network is shown. By this method, the frequency distribution functions are obtained with respect to a network composed of two layers, the network used in information retrieval and the network generating a brain wave.

KEY WORDS: Neural network; frequency distribution function; partition function; information retrieval; brain wave; reverberation.

1. INTRODUCTION

One method of studying information processing in neural networks is by neural dynamics, in which the motion of signals is described by various types of neuronic equations. One of these, the digital equation, describes the time development of the excitation states of cell bodies by terms which take the value one or zero according to whether the cell body is exciting or not. The oscillatory behavior of neural networks has been analyzed in terms of such equations.

The rhythmic excitation of neurons is connected with the behavior of animals; for instance, the oscillation of a neuron ring was regarded as a model for a starfish in a study by means of computer simulation.(1)

---

[1] Department of Applied Physics, Waseda University, Tokyo, Japan.

[2] Present address: Department of Communications, Toukai University, Hiratsuka-shi, Kanagawa-pref., Japan.

The long-lived oscillation of a neuronal excitation—the so-called reverberation arising when a stimulus is given to one or more cell bodies—is found in the network of an isolated cortical slab(2) and in the circuit between the nucleus of the cerebellum and pons.(3) On the other hand, networks undergoing reverberations can be studied in terms of neuronic equations; it has been shown, by means of computer simulation,(4) that reverberation occurs in networks with random coupling constants.

The study of oscillatory behavior in networks should prove to play an important role in increasing understanding of certain physiological phenomena whose exact mechanisms have not yet been confirmed in detail. A brain wave is one such phenomenon. The network generating a brain wave is considered to be in the thalamus or neocortex. However, there have been few studies of the neurodynamical aspects of brain waves. Network models able to account for the difference in the frequency of an electroencephalogram for awake and sleep states should also be able to give information of a general nature on interneuronal coupling.

Another physiological phenomenon of interest here is the maintenance of memory. Some have proposed the idea that reverberation is involved in the process of fixing memory.(5,6) Oscillation, including reverberation, will take place by the circulation of an excitation through closed circuits in a network. This relation between circulation and oscillation is the fundamental idea to be used here in developing the new method of analyzing oscillations, even if the network is too complex to be separated into individual closed circuits.

The localization of oscillation or the possibility of localized oscillation by stimulus is proposed as a model of memory; in a network where the coupling constants take values of one or minus one at random in space, localized oscillation has been studied using a linear differential equation by means of computer simulation.

Oscillations in neural networks are observed in other region besides the brain; for instance, the oscillatory motion of the intestine is due to the oscillation of the excitation of cell bodies in the network.

In the study of oscillation in terms of a digital equation, it needs to be taken into account that noise usually exists in a neural network. Therefore, a stochastic equation is used in the theoretical analysis. The equation for the averaged excitation is called the analog equation in the sense that the output excitation is a continuous function of the input stimulus. An analog neuronic equation is described in general by the form

$$S_i(t + \Delta t) = h\left(\sum_j a_{ij} S_j(t - \tau_{ij}) + \gamma_j\right) \tag{1}$$

where $S_i(t)$ is the excitation probability of the ith cell body at time t, such as a pulse frequency or the ensemble average of the state of the cell body. The nonlinear function $h(x)$ is a monotonic function with respect to x, which theoretical neurologists often approximate by the hyperbolic tangent of x. Equation (1) was found from electrophysiological experiments. Its theoretical derivation from a digital equation has been carried out in the case where $S(t)$ is an averaged pulse frequency when the stimuli are stationary, and when the stimuli vary with time in the form $Ae^{-\alpha t} + B$.(7), [3]

Under the assumption that the excitation probability is not too large, the analog neuronic equation is approximated by the linear form

$$S_i'(t + \Delta t) = \sum_j a_{ij} S_j'(t - \tau_{ij}) \tag{2}$$

where $S_i'(t) = S_i(t) - \epsilon_i$, ϵ_i being the component of the column vector ϵ derived from the unit matrix E, the matrix A with components a_{ij}, and the column vector of components γ_i :

$$\epsilon = (E - A)^{-1} \gamma$$

Equation (2) shows that the signal of value S' is transmitted from the jth cell body to the ith cell body at the transmission rate a_{ij} and that the transmission time is $\tau_{ij} + \Delta t$. The time development of a neural network described by the difference equation is usually obtained by finding the eigenvalues of the matrix A.

In the present article, the author proposes a new method of analyzing the time characteristics of a network, without calculating the eigenvalues in terms of usual matrix algebra. The method is "time partition analysis," and is often useful for arriving at a knowledge of the frequency distribution and the localization of an oscillation; here, the conclusions of the analysis of many-particle systems in terms of the energy partition function developed in statistical mechanics and of the distribution function method applied to high-polymer physics are applied to neural network analysis.

2. TIME PARTITION FUNCTION IN A NETWORK

A network is composed of closed circuits which start from the cell body under consideration (which we call the starting point), pass through other cell bodies, and return to the starting point. We number these circuits and suppose that the transmission rate and the time of circulation in the ith circuit are g_i and τ_i, respectively, and we denote the closed circuit (g_i, τ_i).

---

[3] In the case where $S(t)$ is an ensemble average of the excitation states of neurons, the derivation of Eq. (1) has been carried out for the nonstationary input stimulus by the present author and it has been shown that a_{ij} is the coupling constant in a digital equation divided by the threshold value of a neuron, which plays a role similar to temperature with the dynamical equation of an Ising spin system.

Under the assumption that at the initial time $t = 0$, the excitation probability of a starting point is a definite value and the excitation probabilities of the other cell bodies are zero, the excitation probability of the starting point at time t is the sum of the excitation probabilities for a variety of circulations: The excitation probability after passing through the ith circuit ($i = 1, 2, \ldots, N$) a total of n_i times is multiplied by $\prod_{i=1}^{N} g_i^{n_i}$ and the number of ways for such a circulation is $(\sum_{i=1}^{N} n_i)!/\prod_{i=1}^{N} n_i!$. The excitation probability of the starting point at time t, therefore, is obtained from

$$S(t) = \sum_{\{n_i\}} \left[\left(\sum_i n_i\right)! \Big/ \prod_i n_i!\right] \prod_i g_i^{n_i} \tag{3}$$

where the summation $\sum_{\{n_i\}}$ represents the sum of all the combinations of n_i ($i = 1, 2, \ldots, N$) under the condition that

$$\sum_{i=1}^{N} n_i \tau_i = t \tag{4}$$

We define the average time of circulation through the whole network

$$\langle\tau\rangle = \sum_{i=1}^{N} n_i \tau_i \Big/ \sum_{i=1}^{N} n_i \tag{5}$$

from Eq. (4), this being rewritten as

$$t = \sum_{i=1}^{N} n_i \langle\tau\rangle \tag{6}$$

Let t and $\langle\tau\rangle$ be constant, i.e., from Eqs. (4) and (6), $\sum_{i=1}^{N} (\tau_i - \langle\tau\rangle)\, \delta n_i = 0$ and by Lagrange's method, we obtain the most dominant excitation probability as follows:

$$S_m(t) = S(0)\, e^{\gamma(\alpha) t} \tag{7}$$

where

$$\gamma(\alpha) = \alpha + (1/\langle\tau(\alpha)\rangle) \ln \Xi(\alpha) \tag{8}$$

$$\Xi(\alpha) = \sum_{i=1}^{N} g_i e^{-\alpha\tau_i} \tag{9}$$

and α is determined so as to satisfy

$$\langle\tau\rangle = \langle\tau(\alpha)\rangle = -\partial[\ln \Xi(\alpha)]/\partial\alpha \tag{10}$$

The geometrical expression of Eqs. (8) and (10) is the following: At an arbitrary value of α, the absolute value of the gradient of a tangent of $\ln \Xi(\alpha)$

is the average circulation time and the value of the intersection of the tangent and the α axis is $\gamma(\alpha)$.

With a knowledge of $\Xi(\alpha)$ and by calculating $\langle\tau(\alpha)\rangle$ and $\gamma(\alpha)$ from Eqs. (7) and (8), we can find the time characteristics of a starting point as follows: The excitation probability of a starting point is the wave given as the sum of oscillations with frequencies $1/\langle\tau(\alpha)\rangle$ and amplitudes $e^{\gamma(\alpha)t}$. We therefore picture the modified frequency distribution function γ versus $1/\langle\tau\rangle$; γ has a maximum value at the frequency $1/\langle\tau\rangle$ determined by α of $\Xi(\alpha) = 1$, because $\Xi(\alpha)$ is a decreasing function of α.

3. ANALOGY TO THE DISTRIBUTION OF A POLYMER CHAIN

In high-polymer physics, $G_N(\mathbf{r}, \mathbf{r}')$ represents the distribution function of a high-polymer chain of length N which takes various conformations with one end at point $\mathbf{r}$ and the other at point $\mathbf{r}'$.

The distribution function $G_N(\mathbf{r}, \mathbf{r})$ is analogous to the contribution to the time partition function of a variety of closed circuits which start from the cell body at point $\mathbf{r}$ and return to it after passing through the arbitrary $N - 2$ cell bodies.

We consider a network where the two nearest-neighboring cell bodies are connected by a fiber of length a. In analogy to the distribution function of a high-polymer chain composed of N molecules of equal size a, the contribution to the time partition function $G_N(\mathbf{r}, \mathbf{r})$ [hereafter we write $G_N(\mathbf{r})$] in the network is expressed as

$$G_3(\mathbf{r}) = [\exp -\beta\phi(\mathbf{r})] \int f(\mathbf{r} - \mathbf{r}') \exp -\beta\phi(\mathbf{r}')\, d\mathbf{r}' \tag{11}$$

and, for $N \geq 4$,

$$\begin{aligned} G_N(\mathbf{r}) = \int \cdots \int f(\mathbf{r} - \mathbf{r}_1) f(\mathbf{r}_1 - \mathbf{r}_2) \cdots f(\mathbf{r}_{N-2} - \mathbf{r}) \\ \times \exp -\beta[\phi(\mathbf{r}) + \phi(\mathbf{r}_1) + \cdots + \phi(\mathbf{r}_{N-2})]\, d\mathbf{r}_1\, d\mathbf{r}_2 \cdots d\mathbf{r}_{N-2} \end{aligned} \tag{12}$$

where

$$f(\mathbf{r}) = (1/4\pi a^2)\, \delta(|\mathbf{r}| - a) \tag{13}$$

and, with $\phi(\mathbf{r})$ corresponding to the potential energy at point $\mathbf{r}$ in a high-polymer chain, we define for a network

$$\exp -\beta\phi(\mathbf{r}) = g(\mathbf{r}) \exp -\alpha\tau(\mathbf{r}) \tag{14}$$

under the supposition that the transmission rate and time from the cell body at $\mathbf{r}$ to the one at $\mathbf{r} + \mathbf{R}$ ($|\mathbf{R}| = a$) are given by $g(\mathbf{r})$, independently of the direction of **R**, and $\tau(\mathbf{r})$, respectively. The time partition function is obtained from

$$\Xi(\mathbf{r}) = \sum_{N=3}^{\infty} G_N(\mathbf{r}) \tag{15}$$

The calculation of the distribution function (12) has been carried out by various methods. In this article, we try to calculate (12) in terms of the differential equation. From Eq. (12), we obtain

$$\begin{aligned} G_{N+1}(\mathbf{r}) &= \int f(\mathbf{r} - \mathbf{s})[\exp -\alpha\phi(\mathbf{r})]\, G_N(\mathbf{s}, \mathbf{r})\, d\mathbf{s} \\ &\doteq [\exp -\alpha\phi(\mathbf{r})]\{G_N(\mathbf{r}) + \tfrac{1}{2}a^2\, \nabla_r{}^2 G_N(\mathbf{r}) + \cdots\} \end{aligned} \tag{16}$$

and, for small a, $G_{N+1}(\mathbf{r})$ is approximated as

$$G_{N+1}(\mathbf{r}) = [\exp -\alpha\phi(\mathbf{r})]\{G_N(\mathbf{r}) + \tfrac{1}{2}a^2\, \nabla_r{}^2 G_N(\mathbf{r})\}$$

From Eqs. (11)–(16), we have

$$[-\tfrac{1}{2}a^2\, \nabla_r{}^2 + V(\mathbf{r})]\, \Xi(\mathbf{r}) = Q(\mathbf{r}) \tag{17}$$

$$V(\mathbf{r}) = g^{-1}(\mathbf{r})[\exp \alpha\tau(\mathbf{r})] - 1 \tag{18}$$

and

$$Q(\mathbf{r}) = \frac{1}{4\pi a^2} \int \delta(|\mathbf{R}| - a)\, \frac{1}{1 + V(\mathbf{r} + \mathbf{R})}\, d\mathbf{R} \tag{19}$$

When V has spherical symmetry, let $\theta(r)$ be the solution of the equation

$$[-\tfrac{1}{2}a^2\, \nabla r^2 + V(r)]\, \theta(r) = 0 \tag{20}$$

and we obtain

$$\Xi(r) = \theta(r) \int [1/\theta(r)^2] \left\{ \int^r V(x)\, e^{-2\alpha\tau(x)} \theta(x)\, dx + C_1 \right\} dr + C_2\theta(r) \tag{21}$$

The arbitrary coefficients C_1 and C_2 are decided from the condition that in the power-series expansion of $\Xi(r)$ with respect to $e^{-\alpha\tau(r)}$, the constant term and the linear term of $e^{-\alpha\tau(r)}$ are zero.

For instance, in the case where $\tau(\mathbf{r}) = \tau$ and $g(\mathbf{r}) = g$, we obtain

$$\Xi(\mathbf{r}) = (ge^{-\alpha\tau})^2/(1 - ge^{-\alpha\tau})$$

The modified frequency distribution is shown in Fig. 1.

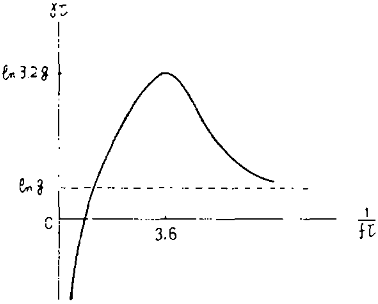

Fig. 1. Modified frequency distribution function. The transmission rate and time between nearest neighbors are g and τ, respectively.

4. DISCRETE FREQUENCY DISTRIBUTION

The frequency distribution of a network is discrete if the differential of the time partion function is discontinuous. In this section, we discuss the discrete frequency distribution of a network, comparing the time partition function with the partition function in the statistical mechanics of a many-particle system.

In a many-particle system when the energy E_i of the particle configuration of the ith state, the fugacity z, and the number of particles n_i are given, the grand partition function and the average density are obtained as [8]

$$\Xi(z) = \sum_i e^{-\beta E_i} z^{n_i} \qquad \text{and} \qquad \rho = z(\partial/\partial z) \ln \Xi(z)$$

where β is the reciprocal of the temperature multiplied by Boltzmann's constant. Then, from the identification of $-n_i \ln z$ and $-\beta E_i$ with $\alpha\tau_i$ and $\ln g_i$, respectively, the time partition function and the average circulation time correspond to the grand partition function and the average density, respectively.

The occurrence of a phase transition at a point of discontinuity of the differential of $\Xi(z)$ with respect to z is compared with the fact that the frequency distribution has a gap in the frequency at the point of the differential discontinuity of $\Xi(\alpha)$ with respect to α. Therefore, we can use the conclusions of a grand partition function analysis for the analysis of the frequency distribution in a network.

For example, we refer to the calculation of a grand partition function in a one-dimensional lattice system with l sites[9] where the nearest-neighbor interaction u_1 is extremely repulsive and the potential u_2 between the second neighbors is attractive if, and only if, there exists a particle between them such that

$$\begin{aligned}\mathscr{E}_1 &= \exp(-\tfrac{1}{2}\beta u_1)\\ &\simeq 0\\ \mathscr{E}_1\mathscr{E}_2 &= \exp(-\tfrac{1}{2}\beta u_1 - \tfrac{1}{2}\beta u_2)\\ &> 0.\end{aligned}$$

In this system, the grand partition function is calculated as

$$\begin{aligned}\Xi(z) &= \sum_{N=0}^{l} \left(\sum_{\Sigma P_i = N} e^{-\beta E} \right) z^N\\ &= \sum_{\{P_i\}} \prod_{i=1}^{l} \mathscr{E}_1^{P_{i-1}P_i} \mathscr{E}_2^{2P_{i-1}P_iP_{i+1}} \mathscr{E}_1^{P_iP_{i+1}} z^{P_i}\end{aligned}$$

where P_i takes the value one or zero according to whether there is or is not a particle at the ith site. Of the variety of sequences $\{P_i\}$, $e^{-\beta E}$ diminishes for those sequences containing the isolated pair of particles, and the contribution of such a sequence of N particles having n isolated particles and cluster of n_k' particles is not zero and is given by

$$\begin{aligned}e^{-\beta E} z^N &= (N!/n!\,k!)(1\cdot z)^n \{\mathscr{E}_1^2(\mathscr{E}_1^2\mathscr{E}_2^2)^{n_1'-2} z^{n_1'}\}\\ &\times \{\mathscr{E}_1^2(\mathscr{E}_1^2\mathscr{E}_2^2)^{n_2'-2} z^{n_2'}\} \cdots \{\mathscr{E}_1^2(\mathscr{E}_1^2\mathscr{E}_2^2)^{n_k'-2} z^{n_k'}\} \qquad (22)\end{aligned}$$

where

$$n + n_1' + n_2' + \cdots + n_k' = N$$

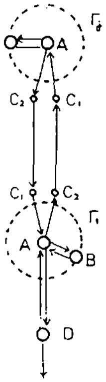

Fig. 2. Network model with discrete frequencies. The direction of signal transmission is indicated by the arrows.

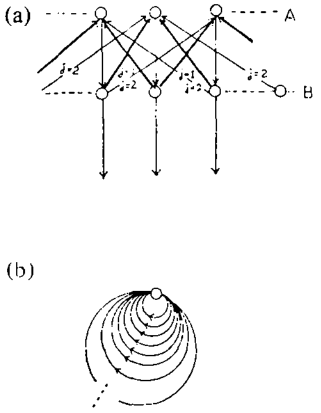

Fig. 3. Network model with discrete frequencies. The direction of signal transmission is indicated by the arrows. Network (a) is topologicaly equivalent to network (b).

and

$$n_i' \geqslant 3, \qquad i = 1, 2, \ldots, k$$

In a network, Eq. (22) represents the contribution of circulation through such closed circuits as $(1, -[1/\alpha]\, n \ln z)$, $(\mathscr{E}_1^2[\mathscr{E}_1^2\mathscr{E}_2^2]^{n_1'-2}, -[1/\alpha]\, n_1' \ln z)$, $(\mathscr{E}_1^2[\mathscr{E}_1^2\mathscr{E}_2^2]^{n_2'-2} - [1/\alpha]\, n_2' \ln z), \ldots$, which are realized in such networks as those illustrated in Fig. 2 or Fig. 3, which are explained in the appendix. The grand partition function of the system for large l is obtained as

$$\Xi(z) = \lambda_{\max}^{l} \tag{23}$$

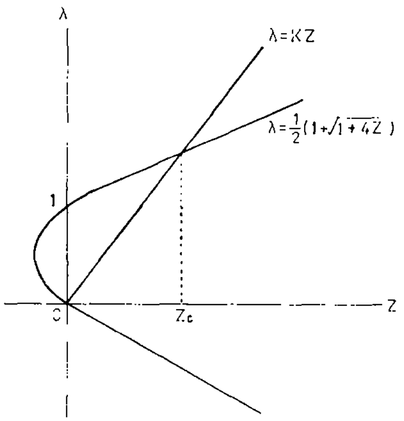

Fig. 4. Variation of λ with z; here, $K = \mathscr{E}_1^2\mathscr{E}_2^2$, $z_c = (K+1)/K^2$.(9)

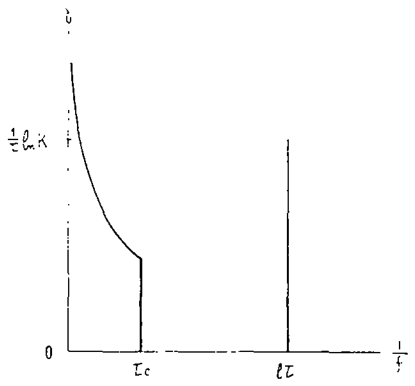

Fig. 5. Modified distribution function with frequency gap; here, $K = \mathscr{E}_1^2\mathscr{E}_2^2$, $\tau_c = l\tau/(K + 2)$. The value of τ_c decreases as K increases.

where $\lambda_{\max}$ is the maximum value of λ, which varies with z as shown in Fig. 4. The system undergoes a phase transition at the point $z = z_c$. Using this conclusion and supposing that $-(1/\alpha)\ln z$ does not depend on n, n_1', n_2',... and is equal to the unit time τ, we can obtain the average of a period of circulation as

$$\langle\tau\rangle = \tau z\, \partial[\ln \Xi(z)]/\partial z \tag{24}$$

The frequency distribution function is illustrated in Fig. 5. The frequency distribution has a gap.

5. SUMMARY AND DISCUSSION

The partition function method was applied to the analysis of the time characteristics of a network in the brain under the condition that the probability of excitation is transmitted in a network in a linear form. When the excitation probability increases, the condition of linearity does not hold, because the output excitation probability will saturate for a large value of input excitation probability. The transmission rates g_i ($i = 1, 2,...$) become small for a large excitation, as shown in Fig. 6, the excitation probability approaching the equilibrium at the g_i which makes the maximum value of $\gamma(\alpha)$ zero for α such that $\Xi(\alpha) = 1$. When the excitation probability decreases for a constant value of g_i, the condition of linearity holds as long as the excitation probability is not extremely small, and then, as the transmission rates become small, the excitation probability decreases more rapidly.

The examples given in the present paper suggest the general types of networks in the brain. The example mentioned in Section 3 is a homogeneous

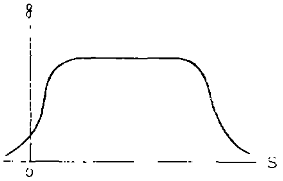

Fig. 6. Transmission rate g versus input stimulus S. The input–output relation is assumed to be $S_{\text{out}} = \tanh(\beta S)$. The transmission rate is defined by $g = dS_{\text{out}}/dS$.

network where the nearest-neighbor cell bodies are connected by the transmission rate g (this is referred to as model I). The transmission rate is a function of the threshold of a cell body and will be inversely proportional to the threshold. As shown in Fig. 1, the frequency of the most dominant oscillation becomes high when the threshold becomes large—this can be regarded as the general property of a network with excitatory connections. As the large-scale networks in the brain, we consider the network where a body Γ_i with excitatory connections g between the cell bodies in it gives rise to an inhibitory effect through the inhibitory cell bodies to the nearest-neighbor bodies Γ_j ($j \neq i, j = 1, 2,...$) in addition to an excitatory effect through the excitatory connection (this is referred to as model II). For large g, the Γ_i are excited and the nearest-neighbor bodies Γ_j are inhibited as shown in Fig. 7, and the most dominant oscillations in the network are those of isolated Γ_i. For small g, the excitation of the Γ_i becomes small and Γ_i no longer inhibits the nearest-neigbor Γ_j and, therefore, the most dominant oscillation becomes that of the circulation through the whole space. The frequency of the oscillation in Γ_i is smaller than that of the whole space because the average circulation time through the whole space is larger than that through Γ_i. This is the model of the structure generating a brain wave. Under the assumption that the transmission rate g becomes small when consciousness transfers from

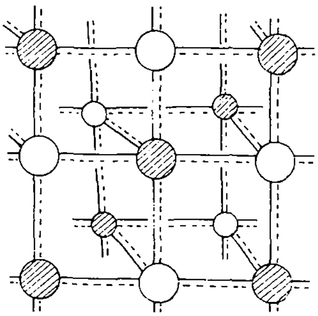

Fig. 7. Symbolic picture of the general connections generating a brain wave. Intrabody connections are excitatory. Interbody connections are inhibitory (dotted lines) and excitatory (solid lines). Inhibited bodies are indicated by shading.

the awake to the sleep state, the model II explains the change of the rhythm from the β rhythm to the α rhythm, whose frequencies are ~10 and ~3–4 Hz, respectively. Therefore, the general type of connection in a large-scale brain network is as follows: the excitatory connections are dominant locally and the inhibitory cell bodies link the local areas, and this is observed over a wide range in the brain.

The network illustrated in Fig. 2 is useful for information retrieval because the circulation with the average time $l\tau$ becomes dominant for large value of g, as shown in Fig. 5, and the signal starting from Γ_i calls on all the bodies at an average period $l\tau$ with a message, e.g., pulse–pulse correlation, average of first passage times, etc.

APPENDIX

We consider the case where $-(1/\alpha)\ln z$ is equal to τ independent of the n_i'.

Figure 2 shows the network composed of a large number of bodies Γ_i $(i = 1, 2, \ldots)$ in which the cell bodies A and B are connected with transmission rate equal to one and transmission time equal to $\tau/2$. Further, a body Γ_i is connected to other bodies Γ_j $(j = 2, 3, \ldots)$ through the "transmission cell bodies" C_1 and C_2 which have transmission rates and times $\mathscr{E}_1$ and τ and $(\mathscr{E}_1\mathscr{E}_2)^{j-1}$ and $(j-1)\,\tau/2$ $(j = 2, 3, \ldots)$, respectively. When the temporal behavior of the network is transmitted through the cell body D to the other regions in the brain, the time characteristics of the area including D and Γ_i $(i = 1, 2, \ldots)$ are as analyzed in Section 4.

The network illustrated in Fig. 3(a) is composed of two layers A and B. The two cell bodies of A and B facing each other are connected with transmission rate equal to one and transmission time equal to $\tau/2$, and, further, there are backward the transmission rates $\mathscr{E}_1^2(\mathscr{E}_1^2\mathscr{E}_2^2)^j$ and the corresponding transmission times $(j + 3/2)\,\tau$ $(j = 1, 2, \ldots)$. Under the supposition that the network is homogeneous and all the cell bodies of A are excited with the same period, the network is topologically equivalent to the network with the closed circuits $(1, \tau)$, $(\mathscr{E}_1^2[\mathscr{E}_1^2\mathscr{E}_2^2], 5\tau/2)$, $(\mathscr{E}_1^2[\mathscr{E}_1^2\mathscr{E}_2^2]^2, 7\tau/2), \ldots$, as shown in Fig. 3(b). When the excitation of B is transmitted out through the lines coming from each cell body of B, the time characteristics of the network are as obtained in Section 4.

REFERENCES

1. R. Suzuki, I. Katsuno, and K. Matano, *Kybernetik* **1**:39 (1971).
2. B. D. Burns, *The Mammalian Cerebral Cortex*, Edward Arnold, London (1958).
3. N. Tsukahara and T. Bando, *Brain Research* **10**:295 (1970).

4. B. G. Farley, *Self-Organizing Systems 1962*, Ed. by M. C. Yovits *et al.*, Spartan Books, Washington, D. C. (1962), p. 535.
5. E. R. Hilgard and D. G. Marquis, *Conditioning and Learning*, Appleton (1940).
6. D. O. Hebb, *The Organization of Behavior*, Wiley, New York (1949).
7. E. R. Caianiello and A. De Luca, *Kybernetik* 3:1 (1966).
8. Terrell L. Hill, *An Introduction to Statistical Thermodynamics*, Addison-Wesley (1960), §§1, 7.
9. N. Saitô and M. Oonuki, *J. Phys. Soc. Japan* **29**:69 (1970).