Hypotheses on the functional roles of chaotic transitory dynamics

Ichiro Tsuda
Research Institute for Electronic Science (RIES); Department of Mathematics, Graduate School of Science; Research Center for Integrative Mathematics, Hokkaido University, Sapporo, Hokkaido, Japan

(Received 16 December 2008; accepted 7 January 2009; published online 31 March 2009)

In contrast to the conventional static view of the brain, recent experimental data show that an alternative view is necessary for an appropriate interpretation of its function. Some selected problems concerning the cortical transitory dynamics are discussed. For the first time, we propose five scenarios for the appearance of chaotic itinerancy, which provides typical transitory dynamics. Second, we describe the transitory behaviors that have been observed in human and animal brains. Finally, we propose nine hypotheses on the functional roles of such dynamics, focusing on the dynamics embedded in data and the dynamical interpretation of brain activity within the framework of cerebral hermeneutics. © 2009 American Institute of Physics. [DOI: 10.1063/1.3076393]

The recent development of measurement techniques in neuroscience has brought about many findings about spatio-temporal dynamics of neural activity. These dynamics have been observed as, among other types of phenomena, a coincidence of random spikes, as coherent activity in neuron assemblies, as the nonstationary transitions between synchronization and desynchronization of oscillatory field potentials, as chaotic population dynamics, as chaotic interspike intervals, i.e., chaotic fluctuations of membrane potentials, as dynamically switching cortical states. The complex spatio-temporal changes of this mesoscopic-level activity has been observed to be not merely random, but transitory dynamics with some conspicuous features, such as, nonstationary, repetitive, itinerant, and chaotic transitions. Focusing on dynamic aspects of the brain, we have adopted the framework of chaotic dynamical systems to interpret the functions of dynamic neural activity emerging in the brain. First, we propose five scenarios for the appearance of chaotic itinerancy, which provides typical transitory dynamics. Second, based on the concepts of chaotic itinerancy, Milnor attractors, and Cantor coding, we present nine hypotheses on the formation of dynamic memory. These hypotheses may account for dynamic functional processes, such as, episodic memory and the itinerant process of cognition. These hypotheses also clarify the biological significance of the chaotic activity observed in the brain.

I. INTRODUCTION

During this decade, measurement techniques in neuroscience have developed greatly, giving rise to many findings about spatio-temporal dynamics of neural activity. However, research interest still seems to be restricted to the act of assigning a function to some specific areas based on the observed activity of neurons or neural assemblies. Although it is particularly important for clinical purposes to investigate which parts of the brain are responsible for a certain specific function, extracting an embedded dynamic order from the extremely complicated behaviors of neural systems would be much more important for the further development of brain research. Here, we take the standpoint of focusing on the dynamics of neural activity for an appropriate interpretation of the corresponding function, taking into account the spatio-temporal scales, called mesoscopic levels, that Walter Freeman proposed to study.1

What are the spatio-temporal scales necessary for understanding brain dynamics and related functions? The theory of phase transitions in physics tells us that ordered motion emerged at a macroscopic level, and associated emergent properties can be described as a collective motion of molecular behaviors at a microscopic level. Here, the collective motion is described by order parameters, which are decoupled from individual molecular activity at a microscopic level. Time scales are associated with spatial scales, so that motion at a microscopic level is much faster than collective behavior. This kind of theory was first developed within the framework of equilibrium phase transitions and critical phenomena, and has been extended to nonequilibrium systems by using bifurcation theory via, for instance, the slaving mode principle.2 These extended theories can be applicable to neural dynamics.3 In a neighborhood of the critical point, at which the transition begins, a complex nonequilibrium motion appears, even in equilibrium systems. The spatial scale of motion reaches over the entire scale of space, from microscopic to macroscopic levels, where fractal patterns become dominant. After the transition, the time scales can be considered as decoupled into two distinct components, one of which represents a slow motion constituting an order parameter, and the other a fast motion that can be rounded off in an averaging process. The averaged motion may appear as either periodic or chaotic behavior. Does this scenario hold in brain dynamics? It could be true if one focuses on a collective motion derived from the interactions of a large number of elementary activity of neurons. It should, however, be noted that spatial scales and time scales do not necessarily match in brain dynamics. This brings about the possibility that the time-dependent motion appears as an ordered motion at a mesoscopic level, as well, which can be described by the time evolution of order parameters. The time-dependent
Ginzburg–Landau, which is usually abbreviated by TDGL, type of equation provides such a typical description.

A far-from-equilibrium state can be realized by a steady flow of energy or matter. In this sense, the brain can be considered to operate as a far-from-equilibrium system. Therefore, the activity of neural assemblies can be interpreted in terms of the dynamic states deriving from the instability of far-from-equilibrium states. Ordered motion may depend on time. It follows that the motion can be described by a quantity, such as, a density function \( p \), which will be a function of time \( t \), space \( x \), and other physical quantities \( s \). For instance, such a quantity could be the membrane potential of a neuron \( v \), or a calcium concentration \( u \), where these quantities depend on space and time, namely \( p(x,v,t) = p[x,v(x,t),u(x,t),t] \). This gives a mesoscopic description in this far-from-equilibrium system. It should be noted that variables \( v \) and \( u \) may depend not only on macroscopic behaviors but also on the microscopic behavior of various types of macromolecules and even genes via learning processes. Another well-known mesoscopic description is provided by the Navier–Stokes equation describing hydrodynamic flow. In this respect, one of the key problems is how one can obtain hydrodynamic limits for neural assemblies from a collection of point neurons. From these considerations, the following proposition may hold.

**Proposition 1:** The brain dynamics measured by electrode, optical recordings, magnetoecephalogram (MEG), or electroencephalogram (EEG) represents the brain activity at the mesoscopic level.

In this paper, by focusing on a certain typical dynamic behavior of the brain, we assert that the transitory dynamics can provide a mesoscopic-level description, which may lead to the description of cognitive function. For the first time, we propose five scenarios for the appearance of chaotic itinerancy, which provides a typical transitory dynamics in high-dimensional dynamical systems. Concerning the relation between cortical transitory dynamics and its cognitive function, we propose nine hypotheses. The reliability of each hypothesis is judged based on the accumulation of reliable experimental data and on the plausibility of interpretation. According to the level of reliability of each hypothesis, we assign a number of asterisks. The larger the number of asterisks, the more reliable the hypothesis, with the largest number being 3.

II. CHAOTIC ITINERANCY AS A TRANSITORY DYNAMICS AT THE MESOSCOPIC LEVEL

The complex spatio-temporal changes of mesoscopic-level activity have been observed to be not merely random, but transitory dynamics with some conspicuous features, such as, nonstationary, repetitive, and chaotic transitions. Typical phenomena observed in laboratories are chaotic transitions between “quasiattractors”\(^4,5,1,6\) irregular transitions between synchronization and desynchronization of subthreshold dynamics in the cat visual cortex,\(^7\) irregular re-entry of synchronization of phase differences in human EEG,\(^8\) and the task-related propagation of wave packets consisting of \(\gamma\)-waves with around 30–90 Hz oscillations and \(\beta\)-waves with around 10–30 Hz oscillations.\(^9,10\) A common feature of these phenomena is that the transition appears to be “chaotic” and “itinerant.”

Furthermore, experimental evidence on spontaneous cortical activity or ongoing activity has accumulated recently. For instance, Kenet \textit{et al.} showed that ongoing activity contains a set of dynamically switching cortical states in V1.\(^11\) They suggested that dynamically switching cortical states may reflect expectations about the sensory inputs. A theoretical investigation on this V1 activity has been published.\(^11\) It has also been suggested that spontaneous cortical activity appears in accordance with wandering mental process due to the activation of default networks.\(^13\)

Despite these findings concerning “nonstationary” transitions, it is misleading that brain activity has been described as a relaxation process to an equilibrium state. In fact, the brain dynamics seems to act at a mesoscopic level in far-from-equilibrium conditions. Furthermore, another misleading theory in conventional brain theory is the theory based on the description of nonstationary and transitory processes by a geometric attractor. Although a theory for associative memories developed by Kohonen, Anderson, Amari, Hopfield, and others,\(^14,15\) and also a theory of neural networks based on the attractor dynamics developed by Amari, Hirsch, Hopfield, Amit, and others\(^16\) played a decisively important role in clarifying the direction of theoretical studies for cognitive functions of the brain, it is apparently incorrect to use these theories for the transitory phenomena mentioned above. Establishing a theory for the observed transitory dynamics and its related cognitive functions, on the other hand, has attracted attention. In this context, we have proposed a theory that those complex phenomena can be interpreted in terms of chaotic itinerancy,\(^19–23\) which can describe a typical transitory dynamics in high-dimensional dynamical systems.\(^24,27,28\) Rabinovich and co-workers have also studied dynamical systems which account for cortical transient phenomena, based on the experimental data of neuronal dynamics for olfactory information processing in insects.\(^29,30\) They have proposed a heteroclinic linking of saddle points or cycles for representation of the transient motion.

One can describe various dynamical states in far-from-equilibrium systems in terms of the concept of attractors in dynamical systems. The steady state is described by a fixed point, the periodic state by a limit cycle, the quasiperiodic state by a torus, and the irregular state by a strange attractor. The fact that a neural network can yield chaotic behaviors has been pointed out by Freeman,\(^1,4–6\) Sompolinsky,\(^31,32\) Tsuda,\(^33\) Körner,\(^34\) Aihara,\(^35\) Arecchi,\(^36\) and others. The roles for chaos in the brain have also been widely studied by Nicolis,\(^37,38\) Tsuda,\(^3,9,24,40\) Freeman,\(^4,1,6\) Skarda,\(^1\) Kay,\(^9,10\) and recently by many others.

A transitory dynamics cannot be explained by these geometric attractors because the transition should be associated with the instability of such a state itself. We have proposed\(^19–22\) the phenomenological concept of “chaotic itinerancy” as what expresses the chaotic transitions between “attractor ruins,” in a neighborhood of which the dynamical orbits experience stagnant motion. In other words, the orbits
behave as if there exist attractors, in the sense that the orbits of positive measure are attracted to those areas after a certain length of time. However, such an attractor area is not asymptotically stable. In this case, we called it an “attractor ruin,” or sometimes use the term “quasi-attractor.”

Using this term, the transitory process can be expressed as a chaotic transition between attractor ruins. Chaotic itinerancy has been found in many systems. Typical systems include globally coupled maps (GCM), coupled map lattices (CML), networks of neuron maps, coupled differential equations (CDE), delay-differential equations (DDE), and skew product transformations (SPT).

Characteristics of chaotic itinerancy have also been clarified. The distribution of the residence time of the stagnant motion follows a power law or an exponential law. The chaotic transition usually occurs in high-dimensional phase space, but, for the case where chaotic orbits are confined in a “narrow tube”-like structure, the main component of transition can be described by low-dimensional chaos. The Lyapunov spectrum has the following three specific characteristics. (1) Many of the Lyapunov exponents accumulate in a neighborhood of zero. (2) The zero exponents besides the direction of orbit (in the case of flow) show large fluctuations and do not converge. (3) Even the largest exponent fluctuates, and shows extremely slow convergence. In this respect, concepts, such as, partial hyperbolicity, nonhyperbolicity, and normally hyperbolic invariant manifolds, have attracted attention.

III. GEOMETRIC ATTRACTORS AND MILNOR ATTRACTORS

An attractor ruin cannot be expressed as a geometric attractor, because a dynamical mechanism must allow transitions between attractor ruins.

One possible mechanism is provided by the use of a Milnor attractor. A Milnor attractor was defined by John Milnor to extend the attractor concept to allow an $\omega$-limit set as an attractor if it has a positive measure for its basin. In the following, we give definitions for both a geometric attractor and a Milnor attractor.

**Definition 1** (geometric attractor): Let $M$ be a compact smooth manifold. Let $f:M \rightarrow M$ be a continuous map on $M$. A trapping region is defined as a subset $N$ of $M$ that satisfies $f(N) \supset \text{inter}(N)$, where $\text{inter}(N)$ is an interior of $N$. For a trapping region $N$ of $M$, such as, this, $A=\cap_{n=0}^{\infty} f^n(N)$ defines an attracting set. A geometric attractor is a minimal attracting set. In other words, an attracting set satisfying topological transitivity is a geometric attractor, simply called an attractor.

A Milnor attractor is an extension of the concept of attractor, whereby a Milnor attractor contains a geometric attractor.

**Definition 2** (Milnor attractor): Let $M$ be a phase space, and $B$ a set. A basinlike region of $B$ is defined as $\rho(B) = \{x | \omega(x) = B, x \in M\}$, where $\omega(x)$ denotes a $\omega$-limit set of $x$. A Milnor attractor is defined as a set $B$ satisfying the following two conditions:

1. $\mu(\rho(B)) > 0$, where $\mu$ is a measure equivalent to the Lesbeque measure.

2. There is no true subset $B'$ of $B$ such that $\mu(\rho(B') \setminus \rho(B)) = 0$.

The condition for a geometric attractor, by which all orbits in a neighborhood of an attractor should be absorbed to the attractor, is not necessarily demanded by the condition for a Milnor attractor. A positive measure of orbits approaching an attractor is necessary. This implies that there could be an orbit leaving an attractor. Therefore, a geometric attractor is a Milnor attractor, but not vice versa. However, in this paper, we will use the term “Milnor attractor” in its narrow sense, i.e., as an attractor characterized by neutral stability. In other words, a Milnor attractor here possesses a positive measure of both attracting orbits and repelling orbits. In this narrow sense, dynamics in a neighborhood of a Milnor attractor are described by higher-order terms than the linear term. For example, in one-dimensional flow dynamics, an evolution equation in a neighborhood of a Milnor attractor can be described by $dx/dt = ax^2 + o(x^2)$, where $x$ denotes a deviation from a Milnor attractor, $o(x^2)$ indicates smaller terms than $x^2$, and $a$ is a positive constant. It is noted that a similar equation also holds for one-dimensional maps. Therefore, in the case of multiple Milnor attractors, each Milnor attractor is placed on its basin boundary.

IV. POSSIBLE SCENARIOS FOR CHAOTIC ITINERANCY

Is there a mathematical concept that correctly represents an attractor ruin? Previously, we have proposed possible scenarios. Here, we treat this issue as an extension of these previous theories.

**Scenario 1:** A three-tuple (chaotic invariant set, Milnor attractors, riddled basins) yields chaotic transitions between attractor ruins.

It is apparent that any transition from a Milnor attractor is impossible without external perturbations because it is an invariant set. External perturbations can be provided by interactions with other systems, and also by external noise. Here, for the first time, we take the GCM into account, as a typical example showing chaotic itinerancy. Then we show that a Milnor attractor associated with a riddled basin brings about chaotic itinerancy.

A GCM is defined as follows: For a one-dimensional map $g^{(i)}:R \rightarrow R$ for $1 \leq i \leq N$, $G:R^N \rightarrow R^N$, $x_{n+1} = G(x_n)$ is determined by the relation

$$x_{n+1} = (1 - \epsilon)g^{(i)}(x_n) + \frac{\epsilon}{(N-1)} \sum_{j=1}^{N} g^{(j)}(x_n) \quad (1 \leq i \leq N),$$

where $n$ is a discrete time, $i, j$ are indices of the map, and $N$ is the number of individual elementary maps.

Kaneko first investigated the case of $g^{(0)}$ being identical logistic maps that produces chaos, and numerically found chaotic itinerancy. A GCM is invariant under the substitution $s$ of individual elementary maps. In other words, a group action $s$ commutes with a dynamical rule $h$, i.e., $hs = sh$. In this sense, a GCM
can be called a symmetric system. This type of symmetric system has been widely studied by Ashwin, Breakspear, and others.58–60

Proposition 2: In a symmetric dynamical system, \((M, h)\), an invariant set under a group action is also invariant for the development of the dynamical system.55

The proof is straightforward. Let \(A(s)\) be an invariant set under a group action \(s\), i.e., \(A(s) = \{sx = x, x \in M\}\). For \(x \in A(s)\), because \(sx = x\), \(h(sx) = h(x)\) holds. By the assumption of symmetry, \(s(h(x)) = h(sx)\), and therefore \(s(h(x)) = h(x)\) follows. This means that \(h(x)\) is invariant under the group action \(s\). Therefore, \(h(x) \in A(s)\). In other words, \(A(s)\) is also invariant under \(h\).

When \(h\) produces chaos, a synchronization state of all elementary individual maps is realized by a one-dimensional chaotic set, which is invariant under any substitution of elementary individual maps. Therefore, by proposition 2, the all-synchronized state must be invariant under dynamical development, which is denoted here by \(H_1\). In the present GCM, there are many other invariant sets representing partially synchronized states. For example, two different synchronized states construct a two-dimensional invariant subspace \(H_2\). Now, assume that a partially synchronized state is stable and represented by a geometric attractor. If the Lyapunov exponent that is normal to \(H_1\) is positive, \(H_1\) is unstable in a normal direction. There is no contradication here. If the sign of the normal Lyapunov exponent of \(H_1\) changes from positive to negative values via a blowout bifurcation,61 and if this bifurcation is local, then the basin of attraction of chaos representing the all-synchronized state becomes riddled. Therefore, the chaotic invariant set becomes a Milnor attractor. Because we assume that a normal Lyapunov exponent to \(H_2\) remains negative because of the locality of the blowout bifurcation, a similar situation happens in a neighborhood of \(H_2\). If the partially synchronized state is chaotic, its basin of attraction may also become riddled.53

When orbits approach an all-synchronized state along its stable manifolds, the orbits begin to behave chaotically via the influence of the chaotic invariant set. While wandering chaotically in a neighborhood of such a set, the orbits meet repelling orbits. Then the orbits begin to leave a neighborhood of the all-synchronized state. A similar situation can happen in a neighborhood of a partially synchronized state. Therefore, such states look like attractor ruins mentioned above. One realization of such a ruin may be a chaotic attractor.62

Related simpler cases have been described as on-off intermittency,54,63 and as in-out intermittency.64 On-off intermittency is an intermittency such that an invariant set, which may represent the all-synchronized state, is a single attractor, whereas, for in-out intermittency, such an invariant set includes plural attractors and/or repellers. It was pointed out by Ott61 that the riddled basin accompanies on-off intermittency, and by Ashwin65 that, for the case of in-out intermittency, the basin of attraction of a chaotic invariant set can become riddled, but that it is an open set for the basin of attraction of a periodic orbit or a fixed point.

Scenario 2: The interacting fixed point type of Milnor attractors can yield chaotic transitions between tori or local chaotic attractors.

For scenario 1, the fact that an all-synchronized state is chaotic is a cause of chaotic transition between partially synchronized states. It would be interesting to investigate whether such a chaotic transition can occur when an all-synchronized state is represented by a fixed-point attractor in a Milnor’s sense, which is different from the case for the above GCM. We have investigated whether or not a coupled map system under this condition can produce chaotic itinerancy.42 The above scenario does not hold for a set of fixed-point Milnor attractors, because the basin of attraction of the fixed point must be an open set. In fact, we did not observe any itinerant transition from those fixed-point Milnor attractors. What we actually observed was chaotic transitions between tori and chaos yielded by the interactions of fixed-point Milnor attractors, whose transitions were associated with a riddled basin, where chaotic transitions occur via crisis-induced chaos.

Scenario 3: A heterodimensional cycle may produce chaotic itinerancy.

Under a similar symmetry, the saddle connections can be robust, as Guckenheimer and Holmes67 proved. This holds under the condition that the sum of the dimensions of the unstable manifolds of one saddle and the stable manifolds of the other one exceeds the dimension of phase space, provided that the sum of the dimensions and the space dimension can be equal in the case of vector fields. In each invariant subspace of symmetric dynamical systems, we can confirm this condition. In fact, heteroclinic cycles can be realized in some neural systems.29 However, with this kind of stabilization condition only, chaotic transitions cannot be expected. What is a mechanism for allowing chaotic transitions, based on the saddle connections?

For simplicity, here we treat the case of a saddle connection between two saddles. Let us denote the saddles by \(S_1\) and \(S_2\). Now suppose that an unstable manifold of \(S_1\) contacts a stable manifold of \(S_2\). The orbit starting from \(S_1\) may construct a heteroclinic orbit connecting to \(S_1\), but this should not be robust, for the following reason. The fact that the sum of the dimension \(n_1^7\) of the unstable manifold of \(S_1\) and the dimension \(n_2^7\) of the stable manifold of \(S_2\) exceeds the space dimension \(N\), i.e., \(n_1^7 + n_2^7 > N\), indicates that \((N - n_1^7) + (N - n_2^7) > N\). The latter inequality means that the sum of the dimensions of the stable manifold of \(S_1\) and the unstable manifold of \(S_2\) cannot exceed the space dimension. The unstable manifold of \(S_1\) contacts the stable manifold of \(S_2\) via an \(n_1^7 + n_2^7 - N\)-dimensional surface. However, the stable manifold of \(S_1\) and the unstable manifold of \(S_2\) cannot contact each other. Therefore, in each neighborhood of two homoclinic orbits, homoclinic chaos appears, i.e., the Shilnikov phenomenon, if the orbits are not restricted to some additional invariant space that reduces the effective dimensionality. However, this condition cannot lead to a transition, such as, chaotic itinerancy.
Let us further consider a heterodimensional cycle.68,69 A diffeomorphism \( f \) has a heterodimensional cycle associated with two saddles \( S_1 \) and \( S_2 \) of \( f \) if the saddles have different indices, i.e., different dimensions of those unstable manifolds \( n^u_1 \neq n^u_2 \). A co-index 1 cycle is a heterodimensional cycle with \( n^u_1 = n^u_2 \pm 1 \). A heterodimensional cycle is not robust for the above reason. However, Bonatti and Diaz proved the following theorem.

**Theorem 1:** (Bonatti and Diaz) Let \( f \) be a \( C^1 \)-diffeomorphism having a co-index 1 cycle associated with a pair of saddles. Then there are diffeomorphisms arbitrarily \( C^1 \)-close to \( f \) that have robust heterodimensional co-index 1 cycles.

By the following theorem, in a neighborhood of a diffeomorphism with a co-index 1 cycle, chaotic behaviors are expected.

**Theorem 2:** (Bonatti and Diaz) Let \( f \) be a diffeomorphism with a co-index 1 cycle that has real central eigenvalues. Then there are diffeomorphisms arbitrarily \( C^1 \)-close to \( f \) that have strong homoclinic intersections associated with saddle-node or with flips.

In this case, we expect transitory behaviors, such as, chaotic itinerancy because of the presence of heteroclinic intersections and the possibility of the appearance of stagnant motion in a neighborhood of heteroclinic tangency. However, further studies are necessary to confirm this assertion.

Furthermore, one may discuss the relation of the appearance of chaotic itinerancy to heteroclinic cycles. It may be interesting to note the memory capacity of networks of competing neuron groups. Rabinovich *et al.* estimated it at approximately \( e(N-1)! \), where \( N \) is the number of neurons, calculating the possible numbers of heteroclinic cycles.29 On the other hand, to calculate the critical dimensionality of the appearance of chaotic itinerancy, Kaneko50 estimated two factors that are supposed to determine the dimensionality for the chaotic transition. Let \( N' \) be the system’s dimension. Let us assume that the number of states in each dimension is two, taking into account the presence of two stable states separated by a saddle. The number of admissible orbits cyclically connecting the subspaces, using, say heteroclinic cycles, increases in proportion to \( (N'-1)! \), whereas the number of states increases in proportion to \( 2^{N'} \). If the former number exceeds the latter, then all orbits cannot necessarily be assigned to each of the states, hence causing the transitions. In this situation, we expect itinerant motions between states. This critical number is *six* for chaotic itinerancy.50,26 We identify \( N' \) with \( N \). In such a case, one may conclude that the transition via heteroclinic cycles appears when the memory capacity is less than the number of states, whereas chaotic itinerancy appears in the opposite condition.

**Scenario 4:** *A normally hyperbolic invariant manifold (NHIM) can yield chaotic itinerancy.*

A NHIM is an extended saddle of high-dimension such that the normal Lyapunov exponents to an invariant manifold are greater than the tangential ones in such a manifold. Komatsuzaki and Toda insisted that a NHIM provides a mechanism for chaotic itinerancy.52 This is because, in a neighborhood of a NHIM, one can expect stagnant motion, and the motion in a NHIM can be chaotic.

**Scenario 5:** *Milnor attractors associated with fractal basin boundaries may yield noise-induced chaotic itinerancy.*

In a symmetric dynamical system, the appearance of a negative Lyapunov exponent in the direction normal to the chaotic invariant set was essential for the transition, but it has been pointed out that the presence of positive normal Lyapunov exponents still brings about a curious transition phenomenon.70 It is known that fractal basin boundaries separate multiple attractors.70 Feudel *et al.* found a chaotic itinerancy-like phenomenon in a double-rotor system with weak noise.70 In this system, many periodic orbits coexist, together with higher periodic orbits possessing very tiny basins, which may disappear under the influence of noise, leaving only low periodic orbits. Similar behavior was found in the KII model by Kozma and Freeman,71 where, because of fractal basin boundaries, long chaotic transients appear before the system falls into a periodic orbit. Orbits are trapped for some time in the vicinity of periodic attractors, but are eventually kicked out by noise, following which the orbits become chaotic again because of the fractality of the basin boundary. Consequently, chaotic transitions between periodic attractors occur, possessing a statistics of residence time in a neighborhood of periodic orbits. This is noise-induced chaotic itinerancy. Noise-induced chaotic itinerancy can occur even with the fixed-point Milnor attractors. This type of chaotic transition is not purely random, even with the addition of uniform white noise, but rather ordered, being caused by the original topology in a neighborhood of such fixed points.

**V. ON THE ROLE OF STATE TRANSITIONS IN THE BRAIN**

**A. Cortical spontaneous activity**

What are the functional roles of dynamic transitions in the brain? It has been believed that the brain responds to stimuli only when the stimuli are actually presented. However, recent observations with fine precision in both space and time show that this is not necessarily true. Spontaneous activity of the brain has been measured via field potentials and electrocorticogram. The brain changes its activity in the absence of stimuli such that the spontaneously activated pattern or ongoing activity is similar to what would appear if the stimulus were actually presented. Consequently, spontaneous activity of the brain shows continual spontaneous transitions between specific patterns.11–13,8 This finding may indicate that the brain is always in an active idling state, with possible responsive patterns being evocated to enable quick responses to any stimuli.

Other kinds of spontaneous activity have also been observed. Freeman and Zhai observed spontaneous activity of animal and human brains,72 and conducted a data analysis in terms of a random number moderated by refractory periods. They found that the spontaneous activity can be characterized by black noise, whose power spectrum density follows \( 1/f^\alpha \), where \( \alpha \geq 2 \). The appearance of black noise activity means that extremely rare events predominate. Another
suppose that a recurrent network consists of neurons, called pyramidal neurons. The hippocampal CA3 and memory.

All of these findings may be related to the formation of episodic memory, because the neural representation of episodes will be regenerated in both the neocortex and the limbic system during a rehearsal of a response. In the following subsection, we propose hypotheses about the role of the hippocampus, after a brief review of the research on the activity of the hippocampal CA3 and memory.

B. Representation of episodic memory

From clinical studies of H.M. and later those of R.B., it has been clarified that the hippocampus and related areas are responsible for the formation of episodic memory. Among others, the function of the neural network in CA3 has been highlighted since Marr’s study of simple memory. Marr considered that the CA3 network provides a mechanism for the representation of associative memory, because it possesses massive recurrent connections. In fact, a network of excitatory pyramidal neurons with recurrent connections has been observed in the hippocampal CA3 and the neocortex.

Mathematical studies of recurrent networks in relation to associative memory have been developed. Now suppose that a recurrent network consists of excitatory neurons, called pyramidal neurons. The th neuron’s activity at time is denoted by . We assume that each neuron’s activity is updated by the formula

\[ x^{(k)}(t) = f\left( \sum_{j=1}^{n} w_{ij} x_j(t) + \text{other inputs} - \text{threshold} \right), \]

where \( f \) is a sigmoid function. We let \( x_{ij}^{(k)} \) denote the activity of the pyramidal neuron for the th memory pattern. In the conventional model of associative memory, if the Hebbian learning algorithm is adopted, geometric attractors can be formed in phase space, each of which is assumed to represent a memory pattern in the th memory, if the Hebbian learning algorithm is adopted, geometric attractors can be formed in phase space, each of which is assumed to represent a memory pattern . Here, the Hebbian learning algorithm is given by the formula of the network activity for a memory number . The argument of the sigmoid function for a neuron is from to . We assume the existence of multiple geometric attractors at the initial stage of the network, each of which represents a memory. By the effect of inhibition, the current state of the network is inhibited, once the current state becomes an attractor, or a slight external noise is added, the network approaches one of the attractors. Once the state becomes an attractor, the current state of the network can no longer be inhibited and recover its stability. Let us assume the existence of multiple geometric attractors can be formed in phase space, each of which is assumed to represent a memory pattern . Here, the Hebbian learning algorithm is given by the formula of the network activity for a memory number . The argument of the sigmoid function for a neuron is from to . We assume the existence of multiple geometric attractors at the initial stage of the network, each of which represents a memory. By the effect of inhibition, the current state of the network is inhibited, once the current state becomes an attractor, or a slight external noise is added, the network approaches one of the attractors. Once the state becomes an attractor, the current state of the network can no longer be inhibited and recover its stability. Let us assume the existence of multiple geometric attractors at the initial stage of the network, each of which represents a memory. By the effect of inhibition, the current state of the network is inhibited, once the current state becomes an attractor, or a slight external noise is added, the network approaches one of the attractors. Once the state becomes an attractor, the current state of the network can no longer be inhibited and recover its stability. Let us assume the existence of multiple geometric attractors at the initial stage of the network, each of which represents a memory. By the effect of inhibition, the current state of the network is inhibited, once the current state becomes an attractor, or a slight external noise is added, the network approaches one of the attractors. Once the state becomes an attractor, the current state of the network can no longer be inhibited and recover its stability. Let us assume the existence of multiple geometric attractors at the initial stage of the network, each of which represents a memory. By the effect of inhibition, the current state of the network is inhibited, once the current state becomes an attractor, or a slight external noise is added, the network approaches one of the attractors. Once the state becomes an attractor, the current state of the network can no longer be inhibited and recover its stability. Let us assume the existence of multiple geometric attractors at the initial stage of the network, each of which represents a memory. By the effect of inhibition, the current state of the network is inhibited, once the current state becomes an attractor, or a slight external noise is added, the network approaches one of the attractors. Once the state becomes an attractor, the current state of the network can no longer be inhibited and recover its stability. Let us assume the existence of multiple geometric attractors at the initial stage of the network, each of which represents a memory. By the effect of inhibition, the current state of the network is inhibited, once the current state becomes an attractor, or a slight external noise is added, the network approaches one of the attractors. Once the state becomes an attractor, the current state of the network can no longer be inhibited and recover its stability. Let us assume the existence of multiple geometric attractors at the initial stage of the network, each of which represents a memory. By the effect of inhibition, the current state of the network is inhibited, once the current state becomes an attractor, or a slight external noise is added, the network approaches one of the attractors. Once the state becomes an attractor, the current state of the network can no longer be inhibited and recover its stability. Let us assume the existence of multiple geometric attractors at the initial stage of the network, each of which represents a memory. By the effect of inhibition, the current state of the network is inhibited, once the current state becomes an attractor, or a slight external noise is added, the network approaches one of the attractors. Once the state becomes an attractor, the current state of the network can no longer be inhibited and recover its stability. Let us assume the existence of multiple geometric attractors at the initial stage of the network, each of which represents a memory. By the effect of inhibition, the current state of the network is inhibited, once the current state becomes an attractor, or a slight external noise is added, the network approaches one of the attractors. Once the state becomes an attractor, the current state of the network can no longer be inhibited and recover its stability. Let us assume the existence of multiple geometric attractors at the initial stage of the network, each of which represents a memory. By the effect of inhibition, the current state of the network is inhibited, once the current state becomes an attractor, or a slight external noise is added, the network approaches one of the attractors. Once the state becomes an attractor, the current state of the network can no longer be inhibited and recover its stability.
How similar are the network structures of the neocortex and the hippocampal CA3? This may stem from biological evolution. The large-cell layers of reptiles developed into the pyramidal cell layers of both the neocortex and the hippocampus in mammals, and the small-cell layers of reptiles developed to the granule cell layers of the dentate gyrus in mammals. The main difference between the neocortex and the hippocampal CA3 comes from the presence of disinhibitory connections from the septal inhibitory neurons to the inhibitory neurons in CA3. This is known as a disinhibitory circuit. Disinhibitions from the septum to the hippocampus are input almost periodically, being synchronized with \( \theta \)-rhythms. On the other hand, \( \theta \)-rhythms are often observed in the hippocampus of animals during searching tasks. Therefore, during a recall of episodes, in the hippocampal CA3, the states of geometric attractor and of Milnor attractor can alternate almost periodically, being associated with the periods of \( \theta \)-waves. We have numerically observed that noise-induced chaotic itinerancy based on the appearance of a fixed-point Milnor attractor, in a model of CA3.

We now propose the following nine hypotheses:

Hypothesis 1*

A memory represented by a geometric attractor in a learning process is represented by a Milnor attractor in a retrieval process. This is caused by the participation of inhibitory neurons in the process.

Hypothesis 2**

The linking process of memories is realized by chaotic activity of the network.

Hypothesis 3**

Chaotic itinerancy provides a neural representation of an episode.

However, the following issue should be considered. For the formation of episodic memory, the mechanism for creating the memory of a time series is necessary, but a single CA3 is not sufficient for this mechanism, although CA3 can yield a time series linking Milnor attractors, as mentioned above. From a mathematical point of view, the distance between patterns can naturally be yielded in CA1, each subset of which represents the time series of finite length that is supposed to be produced from CA3. The distance between the time series of (in)finite length can be measured by a Hausdorff distance between subsets of Cantor sets. However, if episodes are expressed as time series of events reproduced by CA3, episodic memory can be encoded by Cantor sets in CA1. Buszaki found that the GABAergic inhibitory neurons in the septum inhibit the GABAergic inhibitory neurons in the hippocampus, in synchronization with \( \theta \)-rhythms, namely oscillations at 5–8 Hz. Based on this finding, we propose the following hypothesis:

Hypothesis 4****

Disinhibition from the septum to the hippocampus brings about the appearance of attractor dynamics in the hippocampal CA3 if the Hebbian learning algorithm is adopted, because the network effectively becomes a recurrent network of excitatory pyramidal neurons. On the other hand, in the period that this disinhibition is cut off, the overall dynamics in CA3 becomes transitory, because of the instability of the memory space by the operation of inherent inhibitory neurons in CA3. Therefore, the association of one memory representation to another can be repeated in an almost cyclic way, with each appearing approximately every 200 ms.

Furthermore, as for the process of addition of new memories, we propose the following hypothesis:

Hypothesis 5*

When a new pattern is learned, attractor dynamics operate, which allows the representation by a geometric attractor for a learned pattern. After the linking process between memories that include a new memory is strengthened, each elementary memory is represented by a Milnor attractor.

How is the output time series of CA3 represented in CA1? Taking into account the fact that a main direct connection from CA3 to CA1 is unidirectional via the Schaffer collaterals of pyramidal neurons in CA3 and the fact that CA1 dynamics can be contractive, we may suppose that an overall activity of CA1 obeys contractive dynamics driven by chaotic dynamics. We have investigated an abstract CA1 model and derived the Cantor coding. At the next stage of the study, to investigate the biological plausibility of this idea of Cantor coding, we investigated a biology-oriented model that represented the physiological neural networks of CA3 and CA1. To model a single neuron, we used the two-compartment model proposed by Pinsky and Rinzel, which produces quite similar dynamics to the membrane potentials of an actual hippocampal neuron. In both subthreshold and superthreshold dynamics, we found Cantor sets in the membrane potentials of the model CA1 pyramidal neurons. In the biology-oriented model, we found a set of contractive affine transformations, which produces hierarchically ordered patterns of membrane potentials that can be represented by Cantor sets. This means that the dynamics in terms of an iterated function system (IFS) emerges through the network self-organization, thereby producing Cantor sets for the coding of input time series. Furthermore, the membrane potentials for the model CA1 neurons obey a bimodal distribution whose minimum corresponds to the neuron’s threshold. This result may indicate the possibility of decoding the information embedded into Cantor sets by means of a pulse train output of pyramidal cells.

To verify these predictions, we conducted an experiment, using rat hippocampal slices. Random time series of spatial patterns were input to the Shaffer collaterals of pyramidal cells in CA3, with these collaterals making synaptic contacts with pyramidal cells in CA1. We obtained a hierarchical clustering for the membrane potentials of a CA1 neuron, which may indicate the production of Cantor-type patterns in CA1 neurons. We also obtained a return map whenever each elementary pattern in the time series appeared, and then found affine transformations, which appeared to be contractive for most data sets.

The following hypotheses for CA1 dynamics are proposed:
Hypothesis 6***
Episodic memories are encoded in the Cantor sets produced by affine transformations that emerge in the CA1 network.

Hypothesis 7**
Episodic memories are decoded in CA1 outputs as neural pulse trains.

C. Dynamical systems interpretation of the transition between synchronization and desynchronization

Among the data measured at mesoscopic levels, a seemingly random transition between synchronization and desynchronization of neuronal activity is often observed. Some data show the reentry of synchronization at subthreshold activity.7

We made a mathematical model to elucidate a mechanism for this synchronization phenomenon, finding chaotic itinerancy in the process of reentry of synchronization. Let us briefly discuss this issue.

In addition to the class II neuron, which is typically described by the Hodgkin–Huxley (H-H) model, a neuron called a class I neuron is also well known.95 The H-H model is reduced to the two-dimensional phase space, which is known to be a FitzHugh–Nagumo (F-N) model. Recently, many class I neurons have been found also in the higher mammalian brains. In addition to the Na+ and K+ channels adopted in the class II neurons, the class I neurons possess transient potassium channels. By the addition of this channel, the class I neuron has a saddle-node bifurcation as well as a Hopf bifurcation, whereas the class II neuron is typically characterized by a subcritical Hopf bifurcation.

Furthermore, it was recently found that the gap junction-coupled inhibitory neurons are ubiquitous even in the mammalian neocortex. The inhibitory neurons involved in such a network inhibit a pyramidal neuron via their chemical synapses. Therefore, as the first step in the study of these network dynamics, our concern was about the activity of the gap junction-coupled network of class I neurons. To investigate this problem, it is convenient to use a subclass of the class I neurons, called the class I* neurons. A class I* neuron is represented by a vector field on $\mathbb{R}^2$, and is distinguished from others by the following characteristics.28,45

**Definition of class I* neurons
1. The existence of a family of limit cycles possessing a period that becomes infinity at the saddle-node bifurcation point.
2. The existence of a narrow region between two nullclines.
3. The existence of an unstable spiral inside the closed orbit described in 1.

We have studied its gap junction-coupled network.28,44,45

A gap junction can be modeled by nearest-neighbor diffusive couplings.96 For the gap junction-coupled class II neurons, a spiral pattern and a transmission of a pulse front are typical, whereas, for the gap junction-coupled class I* neurons, a transition between synchronization and desynchronization is observed in rather wide parameter ranges. Furthermore, this transition looks chaotic, and therefore it seems to be interpreted as chaotic itinerancy between synchronization and desynchronization. However, detailed investigations led us to a different result. This dynamical system includes a complex structure of Milnor attractors. Dynamical states as components of the transition phenomenon consist of different states, such as, an all-synchronized state, a symmetric metachronal wave, a phase turbulence connecting these two states, and fully developed chaos appearing via a crisis. Here, a metachronal wave means a wave comprising oscillations whose phases monotonically shift in space. An all-synchronized state, a symmetric metachronal wave, and a phase turbulence exist in an invariant subspace $H_M$ with a mirror symmetry. Interestingly, it was numerically shown that a complex of these three states can be a Milnor attractor.

D. Itinerancy of signature and retrieval of memories by an excitatory GABA

One of the striking findings in contemporary brain sciences must be the excitatory GABA. Hiroshi Fujii gave an interesting scenario for the role of an excitatory GABA.97 We provide a brief sketch of this scenario here.

There are two ways that an excitatory GABA may appear. One is that it is excitatory in an early period of postnatal development, and the other is that it is excitatory even for an adult brain. Certain hypotheses have been proposed for its mechanism, but this is still in dispute. As mentioned above, the presence of inhibition in a space of associative memory can trigger an organization of a successive association of memories, and the absence of inhibition or the presence of disinhibition may realize a single association of memory, which then fixes the memory state. Therefore, it can be suggested that a switch between excitation and inhibition of GABA in a certain time scale may give rise to an alternation of stable and unstable dynamics in the retrieval process of memories.

Hypothesis 8**
An irregular and nonstationary transition between synchronization and desynchronization observed in a local field potential of the mammalian brain is interpreted as chaotic itinerancy between fully developed chaos and a Milnor attractor, the latter possibly comprising substates with a certain kind of symmetry.

GABAergic neurons work in the processes of both storage and retrieval of memories in the cerebral neocortex. In the early period after the input of stimuli, during which GABA can be inhibitory, an unstable dynamics is dominant and a successive association of memories therefore occurs via chaotic itinerancy. In the second stage, during which GABA may be excitatory, attractor dynamics can work, and the brain activity therefore converges to a certain memory state. If synaptic plasticity occurs in this second period, then the input will be designated as a memory. On the other hand, if synaptic plasticity occurs at an earlier stage, a chain of associated memories can be designated as an episodic memory.

In relation to this hypothesis, consideration of the large scale of dynamical systems that model the interactions be-
between the hippocampus and the neocortex is crucial for an understanding of the whole scale of the actual memory process. However, this is beyond the scope of the present study.

VI. SUMMARY AND DISCUSSIONS

We have provided five scenarios for the appearance of chaotic itinerancy, which may bring about dynamical interpretations of cortical transitory behaviors. We have also proposed nine hypotheses about the dynamic aspects of memory, in relation to the neural dynamic activity that has been observed in laboratories. Using new techniques of measurement that have recently been developed, and also using many promising ideas discussed in the field of cognitive neuroscience, it is expected that these hypotheses will be established further. In fact, the possibility of Cantor coding in the hippocampal CA1 has been discussed from a neurophysiological point of view, and experimental evidence for it has been reported using the rat hippocampal slice.95

Nonstationary activity of the brain, which is often observed in the laboratories, can be interpreted by itinerant dynamics, for which we suggested chaotic itinerancy as a typical transitory dynamics in high-dimensional dynamical systems. Nonstationary phenomena in the brain are not always described by deterministic dynamical systems. A noisy dynamical system called a random dynamical system may provide an alternative tool to describe such dynamics. Another option for treating the complexity of the brain’s nonstationary activity is its description by stochastic differential equations, and also by partial differential equations. All of these alternatives treat the cortical random activity in infinite-dimensional space. This issue is related to the transitory dynamics treated here, but is not yet properly described in the mathematical framework presented.

ACKNOWLEDGMENTS

The author would like to express his special thanks to other members of the Japanese gang of five, Hiroshi Fujii, Minoru Tsukada, Kazuyuki Aihara, and Shigetoshi Nara, for continual and invaluable discussions about the present subject. He was partially supported by a Grant-in-Aid for Scientific Research, on Priority Areas “Integrative Brain Research.” He was partially supported by a Grant-in-Aid for Scientific Research, on Priority Areas “Integrative Brain Research,” and also by partial differential equations. All of these alternatives treat the cortical random activity in infinite-dimensional space. This issue is related to the transitory dynamics treated here, but is not yet properly described in the mathematical framework presented.

REFERENCES

51. S. Wiggins, Normally Hyperbolic Invariant Manifolds in Dynamical Systems (Springer-Verlag, New York, 1994).
52. Geometric Structure of Phase Space in Multi-Dimensional Chaos: Applications to Chemical Reaction Dynamics in Complex Systems,” Advances


35. D. Albers, private communication.


40. M. Komuro, private communication.


