Dynamical Self-Organization and Formation of Cortical Maps

Jun Zhang

*Neurobiology Group, University of California,
Berkeley, California 94720, U.S.A.*

ABSTRACT

Amari (1989) proposed a general mathematical formulation on the self-organization of synaptic efficacies and response fields under the influence of external stimuli, in which the existence condition and the dynamical stability of cortical maps were obtained for neurons with binary input-output transfer functions. In particular, an integral-differential equation was derived that describes the equilibrium behavior of cortical response fields. Here, we extend their results to neurons with sigmoidal input-output functions. We first examine and derive the existence conditions of the map in the absence of recurrent connections and afferent auto-correlations. Then, by utilizing the mathematically convenient δ-function expansion for well-localized yet otherwise arbitrary recurrent connections, we are able to reduce the master equation into a partial differential equation guaranteed to possess analytic solutions with given boundary conditions (as addressed in the corresponding Cauchy problem). The receptive fields and the excitable regions of the map can thus be studied by standard methods of second-order partial differential equations of elliptic or hyperbolic type.

1 Introduction

The self-organization of the nervous system and the consequential formation of cortical maps have been studied quite extensively (von der Malsburg, 1973; Swindale, 1980; Kohonen, 1982; Linsker, 1986; Miller et al., 1989). Standard techniques include: 1) variations of Hebbian rule for modifying synaptic efficacies; 2) lateral inhibition (in its general sense) for establishing topographical organization of cells as well as sharpening their tuning properties; 3) adiabatic approximation in decoupling the learning phase (on the slow time scale) and the relaxation phase (on the fast time scale) of the network. However, in most cases, only computer simulation results were obtained and therefore provided limited mathematical understanding of the self-organizing neural response fields. In Takeuchi and Amari (1979) and Amari (1989), a general mathematical formulation was presented to study analytically the existence condition, the resolution and magnification properties as well as the dynamical stability of cortical maps. This rather rigorous approach yields very illuminating results. In particular, they suggested by perturbation analysis that, in the presence of periodic boundary conditions of the mapping, the relative values of the afferent spread size and the receptive field size will determine the emergence of a block-like, columnar structure versus a continuous, topographic organization of the map. Since their analysis was restricted to binary-valued neurons only, i.e., neurons with step-function as their input-output (I/O) curve, it is certainly desirable to extend this approach to the more general case of neurons with arbitrary sigmoidal I/O functions.

2 Dynamics of self-organization revisited

Following Takeuchi and Amari (1979), the activity of the cortical neuron at location x (a 2D vector in general) and time t may be described by its net input $u(x,t)$ (post-synaptic membrane potential with respect to the resting state) and output $v(x,t)$ (average firing rate of the spike train) interrelated via some monotone increasing (sigmoidal) input-output function: $v = f(u)$. The cortical neurons in the model receive both thalamo-cortical afferent projections as well as intra-cortical recurrent (lateral) connections. The relaxation of the system is dictated, on the fast time scale t, by the equation

$$\frac{\partial}{\partial t} u(x,y,t,\tau) = -u(x,y,t,\tau) + \int \omega(x,x') f(u(x',y,t,\tau)) \, dx' + \int (s(x,y',\tau) - s_0) \, a(y',y) \, dy' \, ,$$

(1)

where $\omega(x,x')$ characterizes the weighting of recurrent connections within the cortex from location x' to location x, assumed to be unchanged with time; s_0 represents the non-specific thalamic inhibitory influence; $s(x,y,\tau)$ is the synaptic efficacy from the thalamo-cortical afferent (indexed by y) to the cortical neuron (indexed by x) and $a(y,y')$ the thalamo-cortical afferent activity at y' upon the presentation of a sensory stimulus at y, both varying on a slow time-scale denoted by τ and are thus treated as constants on the fast time-scale t. This "adiabatic approximation" allows Amari (1989) to construct a global Lyapunov function $L(u)$ and prove that, upon the stimulus presentation at y, the value of $L(u)$ monotonously decreases as $u(x,y,t,\tau)$ evolves, on the fast time-scale t, according to Eq.(1) and reaches its "equilibrium" solution $\bar{u}(x,y,\tau,s(\cdot),a(\cdot))$ ($\bar{u}$ is a functional of $s(x,y,\tau)$ and $a(y,y')$, and the bar denotes the equilibrium of the relaxation phase). This establishes a mapping between the external stimulus and the cortex that is described by $\bar{u} = \bar{u}(x,y,\tau)$ at time τ. To study the self-organization of this map, the synaptic efficacy $s(x,y,\tau)$ is assumed to be modifiable, on the slow time-scale τ, according to the following equation of learning (Hebbian rule):

$$\frac{\partial}{\partial \tau} s(x,y,\tau) = -s(x,y,\tau) + \eta \int a(y,y') f(\bar{u}(x,y',\tau)) p(y') \, dy' \, .$$

(2)

Note that, during this learning phase, the stimulus presentations are stochastically independent at each time τ with some prescribed probability distribution $p(y')$. Here we set $p(y') = const$ (and thus let absorbed into the constant η) to indicate the normal developmental course. During the learning phase, synapses get "matured" so that $s(x,y,\tau)$ becomes the time-independent $S(x,y)$:

$$S(x,y) = \eta \int a(y,y') f(U(x,y')) \, dy' \, ,$$

(3)

and $\bar{u}(x,y,\tau)$ becomes the time-independent $U(x,y)$:

$$\begin{aligned}
U(x,y) &= \int \omega(x,x') f(U(x',y)) \, dx' + \int (S(x,y') - s_0) \, a(y,y') \, dy' \\
&= \int \omega(x,x') f(U(x',y)) \, dx' + \int k(y,y') f(U(x,y')) \, dy' - \int s_0 \, a(y,y') \, dy' \, ,
\end{aligned}$$

(4)

where $k(y,y')$ is the auto-correlation of the thalamo-cortical afferents defined as

$$k(y,y') = \eta \int a(y'',y') \, a(y'',y) \, dy'' \, .$$

(5)

3 Existence conditions of the map formation

The master equation (4) developed by Amari and colleagues describes the formation of cortical maps as equilibrium solutions to the dynamical self-organization of the neural system. Yet as a non-linear integral-differential equation, it is in general quite unlikely to derive closed-form solutions for further analysis. By assuming $f(u)$ to be the step function, Amari (1989) examined the existence conditions of the discrete map in the absence of recurrent connections in the cortex and afferent auto-correlations (mutually orthogonal signals). Here we consider the same situation, but with the more general sigmoidal function $f(u)$.

From the two simplifications $\omega(x,x') = 0$ and $a(y,y') = \delta(y - y')$, Eq.(4) becomes

$$U(x,y) = \eta\, f(U(x,y)) - s_0 \;, \qquad (6)$$

whose solution is given graphically by the intersection(s) of the curve $V = f(U)$ and the line $V = (U + s_0)/\eta$ in the (U,V)-coordinates, i.e.,

$$U(x,y) = \lambda_i \;. \qquad (7)$$

Here η and s_0 are constants, and so are λ_is. For the intersections to compose a map, the following conditions must be satisfied: 1) there must be more than one intersection (implying that for every stimulus y there are some x's corresponding to $U(x,y) = \lambda_1$ and other x's corresponding to $U(x,y) = \lambda_2$, etc.) and 2) the resting point $(0, V_0)$ must be one of the intersections (i.e., there must be some regions of x's that would not be affected by any particular stimulus y). Here $V_0 = f(U)|_{U=0}$ is the firing rate at rest ($V_0 > 0$ means spontaneous firing). Considering these points, the existence conditions for the map formation in this special situation can be summarized in Fig. 1 and Table 1. The corresponding ranges of of η and s_0 allowing the map formation may also be obtained.
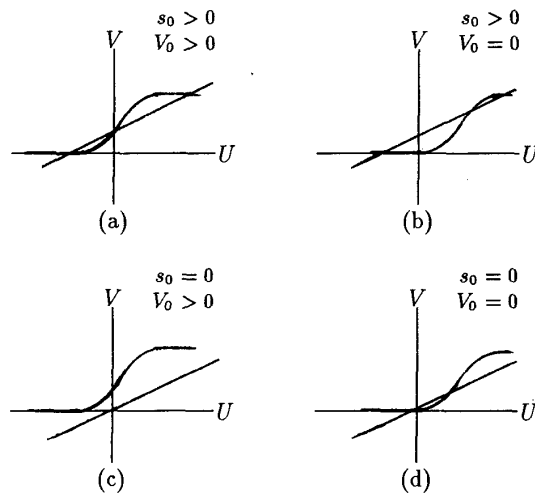

Table 1

Spontaneous Firing Rate / Non-Specific Inhibition	Yes ($V_0 > 0$)	No ($V_0 = 0$)
Yes ($s_0 > 0$)	Map Forms E/I-regions	Not Form
No ($s_0 = 0$)	Not Form	Map Forms E-region only

Fig.1   The straight line and the sigmoidal curve must have more than one intersections, one of which being $(0, V_0)$, in order for the cortical map to form. Therefore, condition (a) or (d) allows map formation, while (b) or (c) doesn't.

4 Excitable region and receptive field of the cortical map

The excitable region in a map refers to the set of cortical locations $\{x : U(x,y) \neq 0\}$ for a particular external stimulus y. The receptive field in a map refers to the set of external stimuli $\{y : U(x,y) \neq 0\}$ at a particular cortical location x. Recurrent connections as well as afferent auto-correlations are both crucial in determining the size of receptive fields and the extend of excitable regions. Amari and colleagues studied the respective resolution and amplification properties, again with the binary-valued I/O function. Here we derive approximations of Eq.(4) for well-localized functions of $\omega(x,y)$ and $a(y,y')$, and meanwhile relax the restriction of $f(u)$ being the step function. Note that in what follows, x, y are taken to be real numbers, i.e., we consider 1D instead of 2D mapping.

We assume that $\omega(x,x') = \omega(|x - x'|)$ (namely translational invariant and symmetric) and take advantage of the following expansion (see Appendix for proof):

$$\omega(x - x') = a_0\,\delta(x - x') + a_1\,\delta'(x - x') + \cdots + a_n\,\delta^{(n)}(x - x') + \cdots ,\tag{8}$$

where the constants a_i are given in Eq.(A.2) of the Appendix. Here $\delta(x)$ is defined by

$$\delta(x) = \begin{cases} \infty & x = 0, \\ 0 & \text{otherwise}, \end{cases}\tag{9}$$

and that

$$g(x_0) = \int_{-\infty}^{\infty} g(x)\,\delta(x - x_0)\,dx .\tag{10}$$

Its derivatives are symbolically defined as

$$\delta^{(n)}(x) = \frac{d^n}{dx^n}\delta(x) ,\tag{11}$$

which clearly have the property (via integration by part)

$$
\begin{aligned}
\int_{-\infty}^{\infty} g(x)\,\frac{d^n\delta(x - x_0)}{dx^n}\,dx &= \int_{-\infty}^{\infty} g(x)\,d\delta^{(n-1)}(x - x_0) \\
&= -\int_{-\infty}^{\infty} \frac{d\,g(x)}{dx}\,\delta^{(n-1)}(x - x_0)\,dx = \cdots\cdots \\
&= (-1)^n \int_{-\infty}^{\infty} \frac{d^n g(x)}{dx^n}\,\delta(x - x_0)\,dx = (-1)^n \left.\frac{d^n g}{dx^n}\right|_{x=x_0} .
\end{aligned}\tag{12}
$$

This, along with Eq.(8), transforms the master equation (4) into the following (we perform the similar expansion for $a(y,y')$, with expansion coefficients b_i):

$$U(x,y) = -s_0 + (a_0 + b_0)\,V - \left(a_1\frac{\partial V}{\partial x} + b_1\frac{\partial V}{\partial y}\right) + \cdots + (-1)^n\left(a_n\frac{\partial^n V}{\partial x^n} + b_n\frac{\partial^n V}{\partial y^n}\right) + \cdots ,\tag{13}$$

where we denote $V(x,y) = f(U(x,y))$. The nature of the δ-function expansion can be envisioned as approximating a well-localized function by successively more "extending" functions $\delta^{(n)}(x)$, while the expansion coefficients a_n and b_n decrease very quickly as n grows larger. (This is particularly evident if we view the δ-function as the limiting case of the normalized Gaussian function whose successive derivatives obviously become less and less localized.) Taking only the leading terms of the expansion, and realizing that $\omega(|x - x'|)$

and $k(|\mathbf{y} - \mathbf{y}'|)$ are both even functions of their arguments, thus making $a_1 = 0$ and $b_1 = 0$, we may truncate the series at the second derivative term (which has the shape of a Mexican hat). Therefore

$$f^{-1}(V(\mathbf{x,y})) = -s_0 + (a_0 + b_0) V + \left(a_2 \frac{\partial^2 V}{\partial \mathbf{x}^2} + b_2 \frac{\partial^2 V}{\partial \mathbf{y}^2} \right) , \tag{14}$$

or

$$g(V) = a_2 \frac{\partial^2 V}{\partial \mathbf{x}^2} + b_2 \frac{\partial^2 V}{\partial \mathbf{y}^2} . \tag{15}$$

Eq.(15) is a semi-linear second-order partial differential equation and, as the corresponding Cauchy problem suggests, always possesses a solution for appropriate boundary conditions (Sneddon, 1957). When $a_2 b_2 > 0$, it is of elliptic type; when $a_2 b_2 < 0$, it is of hyperbolic type. The standard techniques of solving Eq.(15) can be found in mathematical textbooks, such as Sneddon (1957), Chester (1971), and is apparently beyond the scope of this paper.

5 Discussion

The great analytic power of this approach towards a unified description of self-organization of cortical maps, as developed by Amari and colleagues and extended here, overshadows other works of mere computer experimentations that yield little mathematical appreciation of the dynamics as well as the equilibrium behavior of the system. The present formulation embodies the general scheme of layered neural networks with feed-forward (thalamo-cortical) excitations and recurrent (cortico-cortical) connections and takes into account such features like the input auto-correlation in the afferent fibers, and the Hebbian rule of synaptic modification. Future efforts may be directed towards other means of simplifying the master equation (4) to forms easy to handle.

Appendix

We hereby present the proof of the following proposition:
Proposition: Any "well-localized" function $g(\mathbf{x})$ can be expanded into the δ-function and its derivatives

$$g(\mathbf{x}) = a_0 \, \delta(\mathbf{x}) + a_1 \, \delta'(\mathbf{x}) + \cdots + a_n \, \delta^{(n)}(\mathbf{x}) + \cdots , \tag{A.1}$$

where

$$a_n = \frac{(-1)^n}{n!} \int_{-\infty}^{\infty} \mathbf{x}^n \, g(\mathbf{x}) \, d\mathbf{x} . \tag{A.2}$$

By "well-localized", we mean that a_n exist and converge as $n \to \infty$.
Proof: Let the Fourier transform of $g(\mathbf{x})$ be $\hat{g}(\mathbf{k})$, i.e.,

$$\hat{g}(\mathbf{k}) = \int_{-\infty}^{\infty} g(\mathbf{x}) \, e^{i\mathbf{kx}} \, d\mathbf{x} . \tag{A.3}$$

Differentiating $\hat{g}(\mathbf{k})$ n times with respect to $\mathbf{k}$ and then setting $\mathbf{k} = 0$ yield

$$\begin{aligned}
\left. \frac{d^n \hat{g}(\mathbf{k})}{d\mathbf{k}^n} \right|_{\mathbf{k}=0} &= \left. \frac{d^n}{d\mathbf{k}^n} \left(\int_{-\infty}^{\infty} g(\mathbf{x}) \, e^{i\mathbf{kx}} \, d\mathbf{x} \right) \right|_{\mathbf{k}=0} \\
&= \left. \left(\int_{-\infty}^{\infty} g(\mathbf{x}) \, (i\mathbf{x})^n \, e^{i\mathbf{kx}} \, d\mathbf{x} \right) \right|_{\mathbf{k}=0} = i^n \int_{-\infty}^{\infty} \mathbf{x}^n \, g(\mathbf{x}) \, d\mathbf{x} \\
&= (-i)^n \, a_n \, n! .
\end{aligned} \tag{A.4}$$

On the other hand, $\hat{g}(k)$ can be expanded into the Taylor series around $k = 0$:

$$\hat{g}(k) = \hat{g}(0) + \left.\frac{d\,\hat{g}}{dk}\right|_{k=0} k + \cdots + \frac{1}{n!}\left.\frac{d^n\hat{g}}{dk^n}\right|_{k=0} k^n + \cdots . \tag{A.5}$$

This further becomes

$$\hat{g}(k) = a_0 - i\,a_1\,k + \cdots + (-i)^n\,a_n\,k^n + \cdots . \tag{A.6}$$

Therefore, in the inverse Fourier transform of $\hat{g}(k)$

$$
\begin{aligned}
g(x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\hat{g}(k)\,e^{-ikx}\,dk \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}(a_0 - ik\,a_1 + \cdots + (-ik)^n\,a_n + \cdots)\,e^{-ikx}\,dk \\
&= a_0\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-ikx}\,dk\right) + a_1\frac{d}{dx}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-ikx}\,dk\right) + \cdots + a_n\frac{d^n}{dx^n}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-ikx}\,dk\right) + \cdots \\
&= a_0\,\delta(x) + a_1\,\delta'(x) + \cdots + a_n\,\delta^{(n)}(x) + \cdots ,
\end{aligned}
\tag{A.7}
$$

where we have used the identity

$$\delta(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-ikx}\,dx . \tag{A.8}$$

The δ-function and its derivatives $\delta'(x), \delta''(x), \ldots, \delta^{(n)}(x)$ are discussed already in Eqs.(9-12). We thus proved the proposition. $\square$

References

Amari, S. Dynamical study of formation of cortical maps. In Arbib, M.A. and Amari, S. Eds, *Dynamic Interactions in Neural Networks: Models and Data,* pp 15-34. New York: Springer-Verlag, 1989.

Chester, C.R. *Techniques in Partial Differential Equations,* New York: McGraw-Hill, 1971.

Kohonen, T. (1982) Self-organized formation of topologically correct feature maps. *Biol. Cybern.,43,* 59-69.

Linsker, R. (1986) From basic network principles to neural architecture. *Proc. Natl. Acad. Sci. USA,83,* 7508-7512, 8390-8394, 8779-8783.

Malsburg, Ch. von der (1973) Self-organization of orientation sensitive cells in the striata cortex. *Kybernetik 14,* 85-100.

Miller, K.D., Keller, J.B., & Stryker, M.P. (1989) Ocular dominance column development: analysis and simulation. *Science, 245,* 605-615.

Sneddon, I.A. *Elements of Partial Differential Equations,* New York: McGraw-Hill, 1957.

Swindale, N.V. (1980) A model for the formation of ocular dominance stripes. *Proc. R. Soc. Lond. B.,208,* 243-264.

Takeuchi, A., & Amari, S. (1979) Formation of topographic maps and columnar microstructures in nerve fields. *Biol. Cybern.,35,* 63-72.