Mathematical analysis of a neural network with inhibitory coupling

Marie Cottrell

U.A. C.N.R.S. 743 "Statistique Appliquée", Laboratoire de Mathématique, Bât. 425, Université Paris 11, F-91405 Orsay, France, and UFR de Mathématiques, Université Paris 1, F-75634 Paris 13, France

We study the role of inhibition in a nearest-neighbours-connected neural model. The state of the network is a Markov process of which we study the ergodic properties or divergence characteristics using the parameters of the system. We prove that, when inhibition is smaller than a certain threshold, the network is ergodic and works in a stationary way. Conversely, when inhibition increases, the network is divided into two groups: active and inactive neurons. We observe by means of computer simultation that striped or moiré responses appear, whose shape and width depend on considered neighbourhood size. The model resembles the biological reality of the young animal's cerebellar cortex.

AMS 1985 Subject Classifications: 94C99, 92A90.

Markov chains * neural networks * inhibitory coupling * cerebellar cortex * stimuli

1. Introduction

Many formal neural networks have been studied since MacCulloch and Pitts (1943) and Von Neumann (1956). See Kohonen (1984), or Rumelhart (1986) for a review of the different aspects covered in recent research. See also Cottrell and Fort (1986, 1987) or Cottrell (1988a) for some examples of neural networks. In this paper, we study a neural network whose cells are linked by inhibitory coupling.

The model we present here was derived from a simplified model of a neurobiological system, introduced by Axelrad et al. (1985). The motivation was to try to understand the role of the inhibitory synapses which link the main cells (Purkinjie cells) of the cerebellar cortex.

More thorough explanations about the model motivation are set out in Cottrell (1988b) or Axelrad et al. (1987), and neurobiological description of the cerebellar cortex can be found for instance in Eccles et al. (1967), or in Palay and Chan-Palay (1974).

The heuristic description of the model is as follows. The neurons i, $1 \le i \le N$, are arranged in a two-dimensional lattice. Each of them is linked to its neighbours by inhibitory synapses. Let $\mathcal{V}(i)$ be the neighbourhood of the neuron i and $v = |\mathcal{V}|$ the number of its elements. We consider different shapes of $\mathcal{V}$ and different values of

v, which can be 4, 6, 8, etc., as indicated in Figure 1. It is possible to consider the lattice as a tore, so that each cell has the same number v of neighbours.

First, we consider only the spontaneous neuron activity, which consists of electric discharges, so called 'spikes', with constant duration and amplitude. So, we can represent a neuron's activity by a point process, i.e. a sequence of interspike intervals. At time t, the *state* of neuron i is a positive valued variable X_i^t, which represents the waiting time to the next spike of neuron i, i.e. the duration that i should wait if it was not inhibited. Then $X_i^t = 0$ means that i fires a spike at time t.

As long as neither i nor its neighbours fire a spike, then the variable X_i^t decreases linearly with time.

When $X_i^t = 0$, the new i-state is given by a random variable U with distribution $\mathscr{F}$, independent of all the rest of the network.

Each time that a neighbour of i fires a spike, the next i-spike is delayed by a positive parameter θ: this is the inhibitory effect of the synapses between i and its neighbours.

If neuron i is linked to the neurons of $\mathscr{V}(i)$, its trajectory is as pictured in Figure 2.

Due to its simplicity, this model can be considered as a model for any network, neural or not, where only local and inhibitory links exist. It can suggest applications in many other fields: Grossberg's bipole, chemical dissipative structure, crystallography etc.

The paper consists mainly of mathematical analysis of the model, as depending on parameters v, $\mathscr{F}$ and θ.

```
              *                        *
v = 4:   *  ⊙  *        v = 6:   *  *  ⊙  *  *
              *                        *


                                       *
         *  *  *                     *  *  *
v = 8:   *  ⊙  *        v = 12:  *  *  ⊙  *  *
         *  *  *                     *  *  *
                                       *
```
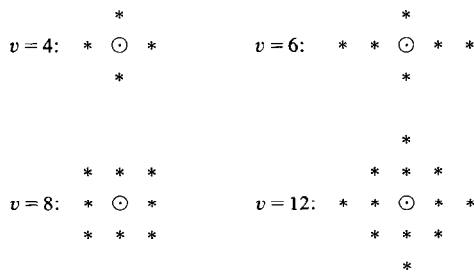
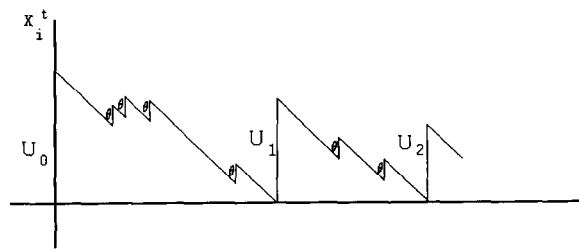Fig. 1. Some neighbourhoods.

Fig. 2. Typical trajectory of a neuron linked to its neighbours.

It is organized as follows: In Section 2, we explicitly define the model, and prove that the network state $(X_i)^t$ is an aperiodic, irreducible, Markov process. Section 3 is devoted to proving that a threshold θ_0 does exist such that, if $\theta < \theta_0$, the $(X_i)^t$ process is ergodic, positive and recurrent, and then its mean interspike interval is evaluated.

Section 4 is devoted to the case $\theta > \theta_0$. For a two neuron system, we prove that at one site the state increases to infinity, while at the other, the state converges weakly to a renewal process with distribution $\mathscr{F}$. In the general case of N cells, stability conditions for different divergent states are given, when $\theta > \theta_0$. In this case, the main result is that in the transient case, the network is divided into two groups:

(1) Active neurons, discharging as if isolated and depending on distribution $\mathscr{F}$.

(2) Inactive completely inhibited neurons.

Computer simulations show that there appear striped or moiré patterns depending on the size v of the considered neighbourhood.

In Section 5, we present some results about the network when it is stimulated by external inputs.

2. The model: Irreducibility, aperiodicity of the Markov process (X^t)

2.1. Equations of the model

The main notations are as introduced in Section 1.

Let $\mathscr{F}$ be the interspike interval distribution of an isolated neuron and let us assume that $\mathscr{F}$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^+$, with density $f > 0$ and that $E(\mathscr{F})$ and $\mathrm{Var}(\mathscr{F})$ are finite-valued.

Let us write X^t for $(X_i)^t$, $1 \leq i \leq N$, and describe the model:

At time $t = 0$, $X^0 = (X_i)^0$ is arbitrarily randomly chosen with positive values.

If at instant t, the network state is $X^t = (X_i)^t$, we have:

Model 2.1.1.

- If no component of X^t is zero, $\mathrm{d}X_i^t = -\mathrm{d}t$ for every i.
- If $X_i^t = 0$ for one site i_0 (i_0 discharges), then:
  - One draws at random an independent random variable U with distribution $\mathscr{F}$, and puts $\mathrm{d}X_{i_0}^t = U - \mathrm{d}t$.
  - For $i \in \mathscr{V}(i_0)$, $\mathrm{d}X_i^t = \theta - \mathrm{d}t$, where θ is the positive delay
  - For $i \notin \mathscr{V}(i_0) \cup \{i_0\}$, $\mathrm{d}X_i^t = -\mathrm{d}t$.

It is clear that the process X^t, defined in its right continuous version, is a continuous-time Markov process.

To avoid the ex aequo, and define only one site i_0, we take as state space

$$\mathscr{S} = (\mathbb{R}^+)^N \cap \{x \mid x_i - x_{i'} \notin \theta\mathbb{Z} \text{ for } i \neq i'\},$$

which allows only one null coordinate.

A trajectory for X^t is typically made up of segments parallel to the diagonal straight line with equation $y_1 = y_2 = \cdots = y_N$ in $(\mathbb{R}^+)^N$, separated by jumps every time they reach a face with equation $y_i = 0$. For instance, if there are only two neurons in the network, one gets a trajectory which is like that in Figure 3.

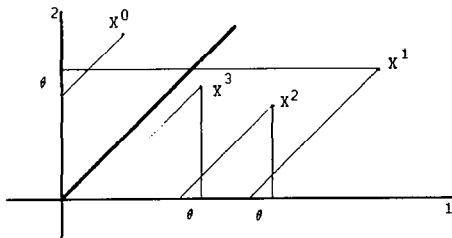Observe that the coordinate processes X_i^t are not Markov processes by themselves.

Fig. 3. Behaviour of X^t in the two-neuron case.

2.2. Embedded chain

With the Markov process defined by Model 2.1.1, we associate the increasing sequence of stopping times $T_1, T_2, T_3, \ldots$, defined as the instants of successive spikes, i.e. the (X^t)'s jump epochs. More precisely, with $T_0 = 0$, we set

$$T_{l+1} = \inf\{t > T_l \,|\, \exists i_0, \, X_{i_0}^{t^-} = 0\}. \tag{2.2.1}$$

Remark 2.2.1. The T_l are all finite a.s., and moreover $\limsup(T_l/l) \leq E(\mathscr{F})$ (see Remark 3.1.1).

For $x \in \mathscr{S}$, let $\mathscr{T}(x) = \min_i(x_i)$ and $i_0(x)$ be the only integer that achieves this minimum.

Then after instant T_l, the neuron number which will be the first to discharge is $i_0(X^{T_l})$, denoted by i^l, and we have

$$T_{l+1} = T_l + \mathscr{T}(X^{T_l}) = T_l + X_{i_0}^{T_l}.$$

Setting $Z(x) = x - \mathscr{T}(x)\mathbb{1}_N$, and defining two vectors with the same $Z(x)$ as equivalent, we have, for $T_l \leq t < T_{l+1}$, $X^t = X^{T_l} - (t - T_l)\mathbb{1}_N$, so that X^t is equivalent to X^{T_l}; therefore $Z(X^t) = Z(X^{T_l}) = Z(X^{T_{l+1}^-}) = X^{T_{l+1}^-}$, denoted by Z^l henceforth.

If we consider the system state just after discharge, at instant T_{l+1}, we get

$$X_i^{T_{l+1}} = Z_i^l + U_{l+1}\mathbb{1}_{(i=i^l)} + \theta\mathbb{1}_{(i \in \mathscr{V}(i^l))} \tag{2.2.2}$$

for $1 \leq i \leq N$ (U_{l+1} is the $(l+1)$th draw with distribution $\mathscr{F}$, corresponding to the $(l+1)$th discharge of a neuron).

The processes X^{T_l} and Z^l are both discrete-time Markov chains. For the chain Z^l, the state space is $\mathscr{S}_0 = \{x \in \mathscr{S} \,|\, \exists i, \, x_i = 0\}$. Its transition kernel is called Q. As for X^{T_l}, its state space is $\mathscr{S}_p = \{x \in \mathscr{S} \,|\, \exists i, \, \forall j \in \mathscr{V}(i), \, x_j > \theta\}$, since every neighbour of the just-fired neuron is simultaneously raised by θ, at each discharge.

Of course, the embedded chain X^{T_i} contains all the information of the process X^t, since the trajectory section going from X^{T_i} to $X^{T_{i+1}}$ is deterministic.

If P is the transition kernel of the homogeneous Markov chain X^{T_i}, we denote its iterates by $P^2, P^3, \ldots, P^m, \ldots$. For an initial distribution μ, we call $\mathbb{P}_\mu$ (or $\mathbb{P}$) the distribution of both processes X^{T_i} and X^t. In a similar way, we put $\mathbb{P}_\mu^k$ for the distribution of X^{T_k}, with initial distribution μ.

2.3. Irreducibility, aperiodicity

Let $Z^0 = X^0 - \mathcal{T}(X^0)\mathbb{1}_N$. The first jump occurs at time $T_1 = \mathcal{T}(X^0)$ and in the site $i_1 = i_0(X^0)$ such that $Z_{i_0}^0 = 0$. We have (2.2.2), with obvious notations,

$$X^{T_1} = Z^0 + U_1 \mathbb{1}_{i_1} + \theta \mathbb{1}_{V(i_1)}.$$

Let $\mathcal{M}(x) = \max_i(x_i) - \mathcal{T}(x)$, then we see that for all $j \neq i_1$, $X_j^{T_1} \leq \mathcal{M}(Z^0) + \theta$, and it is sufficient to choose $U_1 > \mathcal{M}(Z^0) + \theta$, to get that the next jump occurs in $i_1 \neq i_2$.

Then at time T_2, we get

$$X_{i_1}^{T_2} \geq U_1 - \mathcal{M}(Z^0) - \theta, \qquad X_{i_2}^{T_2} = U_2,$$

$$X_j^{T_2} \leq \mathcal{M}(Z^0) + 2\theta, \quad \text{and so on} \ldots.$$

Thus, if at each draw, the innovation U_j is so that

$$U_j > N\mathcal{M}(Z^0) + (1 + 2 + \cdots + N)\theta = N\mathcal{M}(Z^0) + \tfrac{1}{2}N(N+1)\theta,$$

we can define a permutation σ of the integers $\{1, \ldots, N\}$, depending only on Z^0 and defining the order of the sites where there is a jump. Then at time T_N, each coordinate $X_i^{T_N}$ is written $U_{\sigma(i)} + \varphi_i(Z^0)$. And $\varphi_i(Z^0)$ is deterministic, it does not depend on the U_i, and is fully determined by Z^0.

Let us denote by U_σ, $\varphi(Z^0)$ and $(U > a)$ resp. vector $(U_{\sigma(i)})$, vector $(\varphi_i(Z^0))$ and $(U_i > a)$ for all i.

We can sum up this by the following:

Lemma 2.3.1 (Deterministic evolution). *Let $M(Z^0) = N\mathcal{M}(Z^0) + \tfrac{1}{2}N(N+1)\theta$. There exist a deterministic permutation $\sigma(Z^0)$ of $\{1, \ldots, N\}$ and a deterministic vector $\varphi(Z^0)$ with $\max_i(\varphi_i(Z^0)) \leq M(Z^0)$ such that, if $U = (U_1, U_2, \ldots, U_N) > M(Z^0)$, then $X^{T_N} = U_\sigma + \varphi(Z^0)$, with $T_N \leq M(Z^0)$. We also have*

$$T_{N+1} - T_N \leq M(Z^0) + \max_i(U_i). \qquad \square$$

As $\mathcal{S}_0 = \oplus \mathcal{S} \cap \{x_j = 0\}$, by considering in $\mathcal{S} \cap \{x_j = 0\}$ the Lebesgue measure on coordinates different from j and the Dirac measure δ_0 on the jth coordinate, we define (as a direct sum) the measure Λ on $\mathcal{S}_0$.

Lemma 2.3.2 (Lower bound). *If K is a compact set in $\bar{\mathcal{S}}_0$, there is a constant $\gamma_K > 0$, such that*

$$Q^N(z, dz') \geq \gamma_K \mathbb{1}_K(z) \mathbb{1}_K(z') \Lambda(dz').$$

Proof. We need prove it only for $z_1' = 0$. Let $z' = (0, z_2', \ldots, z_N')$. According to the previous lemma, after N steps, the process hits $(U_1 + \varphi_1(Z), \ldots, U_N + \varphi_N(Z))$ and we can choose the U_i, as long as they satisfy $U_i > M(Z^0)$. So, it is possible to give a lower bound of $Q^N(z, dz')$, namely

$$\mathbb{1}_{(z_1'=0)} \Lambda(dz') \int_A^B \prod_{j=1}^N f(u + z_j' - \varphi_j(z) + \varphi_1(z)) \, du \qquad (2.3.1)$$

by choosing A and B that achieve the conditions of Lemma 2.3.1.

As z belongs to the compact K, we can find an upper bound $M(K)$ for $M(z)$. Note that $U_{\sigma(j)}$ must be equal to $z_{\sigma(j)}' + U_1 - \varphi_{\sigma(j)}(z) + \varphi_1(z)$, to 'go down' to z'. Moreover, as $|\varphi_{\sigma(j)}(z)|$, $|\varphi_1(z)| < M(z) \leq M(K)$, it is sufficient to take $B > A > 3M(K)$, in order to get $U_{\sigma(j)} > M(K)$ for all j (lemma's conditions) and positive arguments in the integral expression (2.3.1).

Even if it means to substitute $f \wedge 1$ for f, we may assume (in the lower bound (2.3.1)) that f is a strictly positive bounded integrable function. So the lower bound, which is continuous with respect to $(z_j' - \varphi_j(z) + \varphi_1(z))$, is greater than a constant $C_K > 0$. $\square$

Besides, with one more step, we can hit any z', starting at z, in $N + 1$ transitions, instead of N.

From Lemma 2.3.2, it follows that the chain Z^I is irreducible with respect to the measure Λ, and the last remark provides the aperiodicity.

In fact, for any $c \in \,]0, 1[$, and any initial x, the measure $G^c(x, dy)$ defined as $\sum c^n Q^n(x, dy)$ is equivalent to Λ, and we get:

Lemma 2.3.3. *The chain Z^I is irreducible and aperiodic. The chain Z^I is recurrent (i.e. for all x, all E with $\Lambda(E) > 0$, starting at x, it visits E infinitely often) if there exists x and compact K (with $\Lambda(K) > 0$) such that either*

(1) *starting at x, it returns to K i.o., or*

(2) *the potential $G(x, K) = +\infty$. (See for example Revuz (1975).)* $\square$

We can deduce the same properties for the chain X^{T_i}. Let x and x' belong to a compact $K \subset \mathcal{S}_p$. As the conditional distribution after the first jump depends only on $Z(x)$, after N steps, the chain Z^I goes from $z = Z(x)$ to z', with z' leading to $x' \in \mathcal{S}_p$: if $x' = (x_1', \ldots, x_N') \in \mathcal{S}_p$, there is i such that $x_j' > \theta$ for $j \in \mathcal{V}(i)$ and $z' = (z_1', \ldots, z_N')$ can lead to x' as long as $z_i' = 0$, $z_j' = x_j' - \theta$ for $j \in \mathcal{V}(i)$, $z_{j'}' = x_{j'}'$ for $j' \neq i$ and $j' \notin \mathcal{V}(i)$, provided that $U = x_i'$.

Then in $N + 1$ steps, it is possible to go from x to x', with a positive lower bounded probability on K. Of course, like for Z^I, we can make one more step, for hitting x' in $N + 2$ steps. Hence:

Corollary 2.3.4. *The chain X^{T_i} is irreducible and aperiodic on $\mathcal{S}_p$ (with the Lebesgue measure λ). If there is a compact set with infinite potential, it is recurrent in every*

λ-positive set of $\mathscr{S}_p$. Similarly, X^t is irreducible, aperiodic on $\mathscr{S}$ and the recurrence of X^{T_l} involves that of X^t. □

Indeed, as the chain X^{T_l} contains all the information of the process X^t, X^t has the same properties as X^{T_l}, on its support $\mathscr{S}$.

3. Positive recurrence case

Intuitively the recurrence is equivalent to the fact that each neuron discharges in turn, i.e. each integer i $(1 \le i \le N)$ has a positive probability to verify

$$i = i^l = i_0(X^{T_l}) \quad (X_i^{T_l} = \min_j(X_j)^{T_l}).$$

The *positive* recurrence signifies that the 'bascule' occurs after an integrable time. Conversely, when a neuron i does not discharge any more, it never decreases enough to be the minimum: it is overloaded and $X_i^t \to \infty$.

Let us begin with some technical lemmas.

3.1. Recurrence time for state 0 at each site

In this subsection, it is not necessary to assume that $|\mathscr{V}(i)|$ is the same for all i.

Let i be a given site, with state $x = X_i^0$ at $t = 0$, and let us denote by τ its first return-to-0 time. For $t < \tau$,

$$t < t + X_i^t = x + \theta \sum_{j \in \mathscr{V}(i)} \sum_m \mathbf{1}_{(T_{jm} < t)} \tag{3.1.1}$$

where T_{jm} is the mth return-to-0 time of site j. Denote by $N_j(t)$ the number of returns-to-0 at site j before time t,

$$N_j(t) = \sum_m \mathbf{1}_{(T_{jm} \le t)}.$$

In the initial model, we give us only one sequence U_k i.i.d. with distribution $\mathscr{F}$. We could just as easily assign to each site j a sequence $U_m^{(j)}$ i.i.d. with distribution $\mathscr{F}$. These sequences are independent, and reinitialize site j at each return to 0, so that, for $m \ge 1$,

$$T_{j,m+1} - T_{jm} \ge U_m^{(j)}. \tag{3.1.2}$$

Let $S_m^{(j)} = U_1^{(j)} + \cdots + U_m^{(j)}$ and ν_j be the associated counting process:

$$\{\nu_j(t) \ge m + 1\} = \{S_m^{(j)} \le t\}.$$

So (3.1.2) implies $N_j(t) \le \nu_j(t)$.

Remark that ν_j does not depend on the initial distribution, and that

$$\frac{\nu_j(t)}{t} \xrightarrow{\text{a.s.}} \frac{1}{E(\mathscr{F})}.$$

Remark 3.1.1. Let $N(t)$ be the 'total counting process': $(N(t) \geq l+1) = (T_l \leq t)$. As $\mathcal{T}(X^{T_l}) \leq U_l$ at each T_l, it is clear that $N(t) \geq \nu(t)$, where $\nu(t)$ is the counting process associated with the random walk (U_k), so $\liminf(N(t)/t) \geq 1/E(\mathcal{F})$, that is to say $\limsup(T_l/l) \leq E(\mathcal{F})$.

Since $(t < \tau) \Rightarrow (1 \leq x/t + \theta \sum_{j \in \mathcal{V}(i)} (\nu_j(t)/t))$ (by 3.1.1), we see that the event $(\tau = \infty)$ is impossible if $\theta |\mathcal{V}(i)|/E(\mathcal{F}) < 1$, and we conclude by:

Lemma 3.1.1. *If in site i, $\theta|\mathcal{V}(i)| < E(\mathcal{F})$, then, for every initial distribution, the return-to-0 time at site i is finite a.s.* $\square$

Let us prove now that the hypothesis $\theta|\mathcal{V}(i)| < E(\mathcal{F})$ implies the integrability of τ, with an exponential upper bound. Denote by v the number $|\mathcal{V}(i)|$ of neighbours of i, $\mathbb{P}_x$ the process distribution defined by an initial distribution with $X_i^0 = x$, and ν the common distribution of the ν_j. From $\mathbb{P}_x(\tau > t + x) \leq \mathbb{P}(\theta \sum_{j \in \mathcal{V}(i)} \nu_j(t+x) \geq t)$, we get $\mathbb{P}_x(\tau > t + x) \leq v\mathbb{P}(\nu(t+x) \geq t/v\theta)$. If $[z]$ is the integer part of z, it holds

$$\mathbb{P}_x(\tau > t + x) \leq v\mathbb{P}(S_{[t/v\theta]-1} \leq t + x).$$

But the centred random variable $U - E(\mathcal{F})$ is left-bounded and thus satisfies a Cramer inequality: for every $\alpha > 0$, there exists some $\gamma > 0$, such that

$$\mathbb{P}(S_m \leq m(E(\mathcal{F}) - \alpha)) \leq \exp(-m\gamma).$$

Define $\rho > 0$ by $E(\mathcal{F}) = v\theta(1 + \rho)$, and $\alpha = \frac{1}{2}\rho v\theta$, and let γ be the associated number. Then for every $m \geq 2(2 + x/(v\theta))/\rho$, if $(m + 1)v\theta \leq t < (m + 2)v\theta$, it holds

$$\mathbb{P}_x(\tau \geq t + x) \leq v \exp(-m\gamma),$$

then:

Lemma 3.1.2. *If $\theta < E(\mathcal{F})/|\mathcal{V}(i)|$, for any initial distribution μ (with initial value equal to x at site i),*

$$E_\mu(\tau) \leq A + Bx \quad and \quad E_\mu(\tau^2) \leq A + Bx^2$$

where A and B are constants depending only on distribution $\mathcal{F}$ and delay θ. $\square$

From the inequality $N_i(t) \leq 1 + \nu_i(t)$, we get:

Lemma 3.1.3. *If $\theta < E(\mathcal{F})/|\mathcal{V}(i)|$, then*

$$\mathbb{P}\left(N_i(t) \geq \left[\frac{2t}{E(\mathcal{F})} - 1\right]\right) \leq \exp(-\gamma t) \quad so \ long \ as \ t > t_0$$

where t_0 and γ are positive constants depending only on $\mathcal{F}$. $\square$

Throughout the rest of this section, we assume that all neighbourhoods have the same cardinal v, and that $\theta < E(\mathcal{F})/v$. We take $\mu = \mathcal{F}^{\otimes N}$ as initial distribution.

3.2. Positive recurrence

At first we have:

Lemma 3.2.1. *If* $\theta < E(\mathcal{F})/v$, *the probability sequence*

$$\mu_n = \frac{1}{n} \sum_{k=1}^{n} \mathbb{P}_{\mu}^k$$

is tight (i.e. mass-preserving).

From this lemma, we deduce immediately that every adherent probability is invariant, consequently unique, equivalent to Lebesgue measure λ, and that sequence $(\mathbb{P}^k)$ converges to it. Thus, the chain X^{T_l} is positive recurrent.

Proof of Lemma 3.2.1. Fix a site i, and let $T_{i,k}$ be the sequence of return times to 0 at i (for process X^{i^-}) with $T_{i,0} = 0$. If $T_{i,k} \le s < T_{i,k+1}$, $X_i^s \le T_{i,k+1} - s$ then $\int_{T_{i,k}}^{T_{i,k+1}} X_i^s \, ds \le \frac{1}{2}(T_{i,k+1} - T_{i,k})^2$. But, included for $k = 0$, $X_i^{T_{i,k}}$ has distribution $\mathcal{F}$, so that using Lemma 3.1.2 and putting $C = A + BE_{\mu}(\tau^2)$, it holds

$$E_{\mu} \int_0^{T_{i,k}} X_i^s \, ds \le \frac{1}{2} Ck. \tag{3.2.1}$$

Denote by $\mathcal{E}_s$ the variable $\sum_i X_i^s$. If T_l is the lth discharge time for all the network $T_{i,l} \ge T_l$, and

$$E_{\mu} \int_0^{T_l} \mathcal{E}_s \, ds \le \frac{1}{2} CNl. \tag{3.2.2}$$

Let $\mathcal{E}_k = \sum_i X_i^{T_k}$, and recall that $\mathcal{T}_k = \min_i (X_i)^{T_k}$. The chain changes are given by

$$T_{k+1} = T_k + \mathcal{T}_k, \qquad X_i^{T_k + s} = X_i^{T_k} - s \quad \text{for } 0 \le s < \mathcal{T}_k,$$

and thus

$$\int_{T_k}^{T_{k+1}} \mathcal{E}_s \, ds = (\mathcal{E}_k - \tfrac{1}{2} N \mathcal{T}_k) \mathcal{T}_k \ge \tfrac{1}{2} \mathcal{E}_k \mathcal{T}_k.$$

Observe that $\mathcal{E}_{k+1} = \mathcal{E}_k - N\mathcal{T}_k + v\theta + U_{k+1}$, where (U_j) is the $\mathcal{F}$-distributed innovation sequence, so that for $\alpha = \theta/(N+2)$, we have

$$\text{if } \mathcal{T}_k \le \alpha, \quad \mathcal{E}_{k+1} \ge \mathcal{E}_k. \tag{3.2.3}$$

Case 1. Suppose that $\max(\mathcal{T}_k, \mathcal{T}_{k+1}, \dots, \mathcal{T}_{k+N+1}) \ge \alpha$. Previous formula imply

$$\int_{T_k}^{T_{k+N+2}} \mathcal{E}_s \, ds \ge \tfrac{1}{2} \alpha \mathcal{E}_k.$$

But among instants $T_{k+1}, \dots, T_{k+N+2}$, one of the sites discharges at least twice, so that if we put $W_{k+1} = \min\{U_{k+1}, \dots, U_{k+N+2}\}$, $T_{k+N+2} - T_k \ge W_{k+1}$.

Case 2. Now let $\max(\mathcal{T}_k, \ldots, \mathcal{T}_{k+N+1}) < \alpha$. We know that $\mathcal{E}_{k+j} \geq \mathcal{E}_k$ $(1 \leq j \leq N+2)$, and thus $\int_{T_k}^{T_{k+N+2}} \mathcal{E}_s \, \mathrm{d}s \geq \frac{1}{2}\mathcal{E}_k W_{k+1}$. Let $\tilde{\alpha}_{k+1} = \alpha \wedge W_{k+1}$. By combining both lower bounds, it follows that

$$\int_{T_k}^{T_{k+N+2}} \mathcal{E}_s \, \mathrm{d}s \geq \frac{1}{2}\tilde{\alpha}_{k+1}\mathcal{E}_k. \tag{3.2.4}$$

Observe that, by construction, $\tilde{\alpha}_{k+1}$ is independent of the process evolution until T_k, and has mean $2\beta > 0$. Therefore, there exists $\beta > 0$ such that for any initial distribution,

$$E_\mu \int_{T_k}^{T_{k+N+2}} \mathcal{E}_s \, \mathrm{d}s \geq \beta E_\mu \mathcal{E}_k \quad (k \geq 0). \tag{3.2.5}$$

Just as well, for any initial distribution,

$$(N+2)E_\mu \int_{T_1}^{T_{1+(j+1)(N+2)}} \mathcal{E}_s \, \mathrm{d}s \geq \beta \sum_{k=1}^{1+j(N+2)} E_\mu \mathcal{E}_k.$$

From this and (3.2.2), using definition of μ_n and putting $\mathcal{E} = \sum X_i$, we get

$$\int \mathcal{E} \, \mathrm{d}\mu_{1+j(N+2)} \leq \frac{C}{2\beta} \frac{(j+1)(N+2)+1}{1+j(N+2)} N(N+2), \tag{3.2.6}$$

which proves, as C and N are finite and $\beta > 0$, that $\int \mathcal{E} \, \mathrm{d}\mu_n$ is bounded.

If K_T is the compact $(\mathcal{E} \leq T)$, it follows that $T\mu_n(K_T^C)$ is bounded, hence the tightness of (μ_n) is proved. □

Then, as the chain X^{T_i} is recurrent in every λ-positive subset of its state space, we can deduce the same property for process X^t, in every λ-positive subset of its own state space. Of course, it also has an invariant probability equivalent to Lebesgue measure λ. In the same way as in the beginning of this section, it is sufficient to prove that, for the same initial distribution μ, the family

$$\mu_t = \frac{1}{t} \int_0^t \mu P_s \, \mathrm{d}s \tag{3.2.7}$$

is tight.

Let A_t be the event (for any site i, $1 + N_i(t) \leq k_t$) where k_t is the integer $[2t/E(\mathcal{F})]$. Let $Z_t = 1_{A_t} \int_0^t \mathcal{E}_s \, \mathrm{d}s$. In A_t, $T_{i,k_t} \geq t$ and using (3.2.1),

$$E_\mu(Z_t) \leq \frac{1}{2}CNk_t \leq CNt/E[\mathcal{F}]. \tag{3.2.8}$$

From the Markov inequality,

$$\mu_t[\mathcal{E} > T] \leq P_\mu(A_t) + CN/(TE[\mathcal{F}]).$$

But Lemma 3.1.3 implies that, when t tends to infinity, $P_\mu(A_t)$ tends to 0 and (3.2.7) is proved. So we can sum up the results.

Proposition 3.2.2. *If $\theta < E(\mathcal{F})/v$, and $E(\mathcal{F}^2) < \infty$, the processes X^t and X^{T_i} are ergodic, irreducible, aperiodic and positive recurrent.* $\square$

This result can be extended to the no-homogeneous case: $\sup_i |\mathcal{V}(i)| \theta < E(\mathcal{F})$ is the only hypothesis necessary to ensure that all T_{ik} are finite a.s.

3.3. Distribution of the interspike interval in a two-neuron network (recurrent case)

We consider the process (X^t) for a network which consists of two neurons denoted by 1 and 2, and we study the interspike intervals of neuron 2. In that case, X^t is positive recurrent if $\theta < \theta_0 = E(\mathcal{F})$, since $|\mathcal{V}| = 1$.

Suppose that at time t_0^-, the system state is $(x_0, 0)$. We study the conditional distribution of $(X_1, X_2)^t$ given $X^{t_0} = (x_0, 0)$:

$$
\begin{array}{c|c}
t_0^- & t_0^+ \\
\text{State:} \quad (x_0, 0) & (x_0 + \theta, u_0).
\end{array}
$$

If $u_0 < x_0 + \theta$, the next spike of 2 occurs at $t_0 + u_0$:

$$
\begin{array}{c|c}
t_1^- = t_0^- + u_0 & t_1^+ \\
(x_0 + \theta - u_0, 0) & (x_0 - u_0 + 2\theta, u_1).
\end{array}
$$

If not, the next discharge is at site 1 and it occurs at $t_0 + x_0 + \theta$ in $(0, u_0 - x_0 - \theta)$ and $(u_1, u_0 - x_0)$ after the jump.

Recursively, we can see that time τ between two consecutive discharges of neuron 2 can be written

$$\tau = U_0 + K\theta \tag{3.3.1}$$

where U_0 is a random variable with distribution $\mathcal{F}$, and K an integer-valued variable. We get

$$\mathbb{P}(K > 0) = \mathbb{P}(U_0 > x_0 + \theta),$$

$$\vdots$$

$$\mathbb{P}(K > k) = \mathbb{P}(U_0 > x_0 + \theta, U_1 < U_0 - x_0, U_2 < U_0 - U_1 + \theta - x_0, \ldots,$$

$$U_k < U_0 - U_1 - \cdots - U_{k-1} + (k-1)\theta - x_0).$$

If $U' = U - \theta$ $(EU' = EU - \theta)$,

$$\mathbb{P}(K > k) = \mathbb{P}(U_0 > x_0 + \theta, U_1' < U_0' - x_0, U_1' + U_2' < U_0' - x_0, \ldots,$$

$$U_1' + U_2' + \cdots + U_k' < U_0' - x_0)$$

$$= \mathbb{P}(U_0' > x_0, U_1' < U_0' - x_0, \ldots, U_1' + U_2' + \cdots + U_k' < U_0' - x_0).$$

$$\tag{3.3.2}$$

But $U_1' + \cdots + U_k'$ is a random walk such that $EU_i' = EU - \theta = \theta_0 - \theta$ (2 neurons). We find again that $K \to +\infty$ when $EU' < 0$, i.e. when $\theta > \theta_0$.

Conversely, if $\theta < EU = \theta_0$, $EU' > 0$, and

$$\mathbb{P}(K = k) = \mathbb{P}[(U_0' > x_0) \text{ and } k \text{ is the first strictly positive integer such that}$$
$$U_1' + \cdots + U_k' > U_0' - x_0], \qquad (3.3.3)$$

where (U_i') is the random walk $(U_i - \theta)$. (3.3.3) provides the conditional distribution of K given the state $(x_0, 0)$ just before the discharge of neuron 2.

We have

$$E_{x_0}(K) = \sum_k \mathbb{P}(K > k)$$

$$= \sum_k \int_{x_0+\theta}^{+\infty} f(u_0) \, du_0 \int_0^{u_0-x_0} f(u_1) \, du_1$$

$$\times \cdots \times \int_0^{u_0-x_0-(k-2)\theta-u_1-u_2-\cdots-u_{k-2}} f(u_{k-1}) \, du_{k-1}$$

$$\times \int_0^{u_0-x_0-(k-1)\theta-u_1-u_2-\cdots-u_k} f(u_k) \, du_k.$$

When $K = k$, the interval τ takes the value $u_0 + k\theta$, with $k > 0$ when $u_0 > x_0 + \theta$. At $t_{k+1}^- = t_0 + u_0 + k\theta$, the state is $(u_1 + \cdots + u_k - u_0 + x_0 - (k-1)\theta, 0)$ and if the system is stationary, we can write

$$U_1 + \cdots + U_k - U_0 + X_0 - (K-1)\theta \underset{\mathcal{D}}{\sim} X_0,$$

or also

$$U_1' + \cdots + U_k' - U_0' + X_0 \underset{\mathcal{D}}{\sim} X_0.$$

In particular

$$E(U_1' + \cdots + U_k') = EU_0' = \theta_0 - \theta$$

and since K is a stopping time,

$$(EK)(EU') = (EU_0') \quad \text{hence } EK = 1$$

hence:

Proposition 3.3.1 (Two-neuron case). *For the stationary limiting distribution (when $\theta < \theta_0 = E(\mathcal{F})$), the mean interspike interval is $E(\mathcal{F}) + \theta$.* $\square$

This result can be also deduced like in the next subsection.

3.4. Interspike interval in a N-neuron network (recurrent case)

Let τ be the return-to-0 time of site i: $\tau = U + \theta \sum_{j \in \mathcal{V}(i)} N_j(\tau)$, so, with initial distribution $\mu = (\mathcal{F})^{\otimes N}$,

$$T_{i,k} = U_0^i + \cdots + U_{k-1}^i + \theta \sum_{j \in \mathcal{V}(i)} N_j(T_{i,k}).$$

But the ergodicity implies that $T_{i,k}/k$ converges a.s. to a limit s_i $(0 < s_i < \infty)$. As

$$\frac{U_0^i + \cdots + U_{k-1}^i}{k} - E(\mathcal{F}) \xrightarrow{\text{a.s.}} 0 \quad \text{and} \quad N_i(T_{i,k}) = k,$$

it follows

$$E(\mathcal{F})\frac{N_i(T_{i,k})}{T_{i,k}} + \theta \sum_{j \in \mathcal{V}(i)} \frac{N_j(T_{i,k})}{T_{i,k}} \xrightarrow{\text{a.s.}} 1.$$

But $N_i(t)/t$ converges a.s.; we denote by $1/\tau_i$ its limit, so we get limit equations

$$1 = E(\mathcal{F})\frac{1}{\tau_i} + \theta \sum_{j \in \mathcal{V}(i)} \frac{1}{\tau_j} \quad \forall i. \tag{3.4.1}$$

These convergences take place for any initial distribution. Therefore τ_i is the mean return-to-0 time for site i and for the stationary limiting distribution. In particular, if the network is spatially homogeneous ($|\mathcal{V}(i)| = v$ independent of i),

$$\tau_i = E(\mathcal{F}) + \theta v \quad \forall i. \tag{3.4.2}$$

Hence:

Proposition 3.4.1 (Case of N neurons, with homogeneous neighbourhoods with v elements). *For the stationary limiting distribution ($\theta < \theta_0 = E(\mathcal{F})/v$), the mean inter-spike interval is $E(\mathcal{F}) + v\theta$.* $\square$

4. Divergence

4.1. Two-neuron case (divergent case)

Here $\theta > \theta_0 = E(\mathcal{F})$ ($N = 2, |\mathcal{V}| = 1$). Remember that in any case the T_l are finite a.s.

In the two-neuron case, the sign of $(X_2 - X_1)^{T_l}$ indicates what the next discharge will be: neuron 1 (resp. 2) discharges if $(X_2 - X_1)^{T_l} > 0$ (resp. < 0).

At first, consider the discrete-time process $D^l = (X_2 - X_1)^{T_l}$. We have $D^{l+1} = D^l + \operatorname{sgn}(D^l)(\theta - U_{l+1})$, where U_{l+1} is a r.v. with distribution $\mathcal{F}$, independent of D^l. The mean jump of D^l has the same sign as D^l if and only if $E(\mathcal{F}) - \theta < 0$. See Figure 4. From Section 3, we know that if $\theta < \theta_0$, X^{T_l} and X^l are ergodic. It remains to study the behaviour of X^{T_l} and X^l when $\theta > E(\mathcal{F})$.
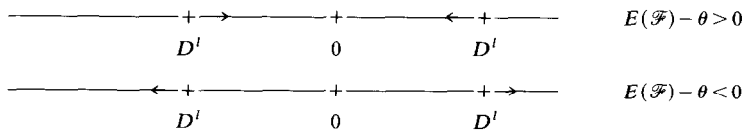

Fig. 4.

Theorem 4.1.1 (Two-neuron case, $\theta > E(\mathcal{F})$). *When $\theta > E(\mathcal{F})$:*
  (i) *$|X_2 - X_1|^{T_i}$ and $|X_2 - X_1|^t$ tend towards $+\infty$ a.s., and more precisely,*
  (ii) *for almost all trajectories ω, when $t \to +\infty$, $(X_2 - X_1)^t$ tends towards $+\infty$ or towards $-\infty$.*

Proof. (i) For $U \sim \mathcal{F}$, we put $V = \theta - U$, $EV = \theta - E(\mathcal{F}) = \alpha > 0$, Var $V = \text{Var}(\mathcal{F})$. We put $D^l = (X_2 - X_1)^{T_l}$. We know that $D^{l+1} = D^l + \text{sgn}(D^l)(\theta - U_{l+1})$ with U_{l+1} independent from D^l, and with distribution $\mathcal{F}$. From $D^{l+1} = D^l + \text{sgn}(D^l) \cdot V_{l+1}$, we get

$$|D^{l+1}| = ||D^l| + V_{l+1}|.$$

$|D^l|$ is D^l-measurable. Let E^l be the conditional expectation given D^l. Then

$$E(|D^{l+1}||D^1, \ldots, D^l) = E^l(|D^{l+1}|)$$

$$\geq |E^l(|D^l| + V_{l+1})| = |D^l| + \alpha > |D^l|.$$

Thus $|D^l|$ is a positive submartingale. As

$$E(|D^{l+1}|) = E(E^l|D^{l+1}|),$$

it holds

$$E(|D^{l+1}|) \geq E(|D^l|) + \alpha.$$

So the sequence $E(|D^l|) \geq l\alpha$, is increasing and goes to $+\infty$.

On the other hand, the submartingale $(|D^l|)$ can be split up (Doob's decomposition) into the sum of an integrable martingale M_l and an increasing process A_l, i.e.

$$|D^l| = M_l + A_l.$$

This decomposition is unique if we take $A_0 = 0$ and $M_0 = |D^0|$. These variables are defined by

$$A_{l+1} - A_l = \Delta A_l = E^l(|D^{l+1}| - |D^l|) = E^l(|D^{l+1}|) - |D^l| \geq \alpha,$$

and $M_{l+1} - M_l = \Delta M_l = |D^{l+1}| - E^l(|D^{l+1}|)$. Thus $A_l \geq l\alpha$ and $A_l \xrightarrow{\text{a.s.}} +\infty$ when $l \to +\infty$.

It remains to study the behaviour of martingale M_l. The submartingale M_l^2 is integrable since distribution $\mathcal{F}$ is square integrable. It can be split up into $M_l^2 = m_l + B_l$, where m_l is an integrable martingale and B_l an increasing process. We have

$$E^l(\Delta M_l^2) = E^l((|D^{l+1}|)^2) - (E^l(|D^{l+1}|))^2$$

$$= E^l(|D^l + \text{sgn}(D^l) V_{l+1}|^2) - (E^l|D^l + \text{sgn}(D^l) V_{l+1}|)^2$$

$$\leq E^l(|D^l + \text{sgn}(D^l) V_{l+1}|^2) - (E^l(D^l + \text{sgn}(D^l) V_{l+1}))^2$$

$$\text{(since } |EX| \leq E|X|)$$

$$= (D^l)^2 + 2(\text{sgn } D^l) \cdot D^l \cdot EV_{l+1} + E[(V_{l+1})^2]$$

$$- (D^l)^2 - 2(\text{sgn } D^l) \cdot D^l \cdot EV_{l+1} - E[(V_{l+1})]^2$$

$$= \text{Var } V = \text{Var}(\mathcal{F}).$$

As

$$B_{l+1} - B_l = E^l(M_{l+1}^2) - M_l^2 = E^l(\Delta M_l^2),$$

we get

$$B_l \leq l \, \text{Var}(\mathcal{F}).$$

Then (see Neveu, 1972, Section VII.2), on $(B_\infty < +\infty)$, M_l converges almost surely to a finite limit; on $(B_\infty = +\infty)$, $M_l/\sqrt{B_l} \ln(B_l) \xrightarrow{\text{a.s.}} 0$ when $l \to +\infty$, and thus

$$M_l = o(\sqrt{l} \ln(l)) \quad \text{a.s.}$$

In both cases, the submartingale $|D|^l$ tends almost surely towards $+\infty$, since $A_l \to +\infty$ a.s. with $A_l \geq l\alpha$ and $M_l/A_l \xrightarrow{\text{a.s.}} 0$. As $(X_2 - X_1)^t$ is equal to D^l for

$$T_l \leq t < T_{l+1}, \qquad |X_2 - X_1|^t \xrightarrow{\text{a.s.}} +\infty \quad \text{when } t \to +\infty.$$

(ii) $|D^l| = A_l(1 + M_l/A_l)$ with $A_l \geq l\alpha$ and $M_l/A_l \to 0$. So there exist β $(0 < \beta < \alpha)$ and $L(\omega)$, such that $\forall l > L(\omega)$, $|D^l| \geq l\beta$.

When $|D^l| \geq l\beta$, a sign change has a probability less than

$$\mathbb{P}(|V_{l+1}| \geq l\beta) \leq \frac{E(V_{l+1}^2)}{l^2\beta^2}$$

which is a convergent series. So according to Borel-Cantelli's lemma, D^l cannot change its sign infinitely often: there is no oscillation.

It is the same for $(X_2 - X_1)^t$. Then for every trajectory, it can be defined a time $T(\omega)$ after which $(X_2 - X_1)^t$ keeps a constant sign and tends either towards $+\infty$, or towards $-\infty$, from (i). $\square$

Let us prove that when one component of (X^t) tends towards $+\infty$, the other converges in law.

Proposition 4.1.2 (Two-neuron case). *When $\theta > \theta_0 = E(\mathcal{F})$, on almost all trajectories, one of the two components tends towards $+\infty$, the other converges in law to the distribution $\mathcal{F}$.*

Proof. On each trajectory, from Theorem 4.1.1, one of the two components tends towards $+\infty$, since the other is positive. Let ω be a trajectory on which $X_1^t \to +\infty$. Let $\tau_A(\omega)$ be the hitting time for $(X_1(\omega) > A)$ and E_k be the event $U_1 + U_2 + \cdots + U_k - k\theta < A - \theta$. As long as $E_1, E_2, \ldots, E_k$ are realized, neuron (2) does discharge. We can write

$$\mathbb{P}\left(\bigcap_{k=\tau_A(\omega)}^{+\infty} E_k\right) > \varphi(A), \tag{4.1.1}$$

where φ is some function of A with $\varphi(A) \to 1$ when $A \to +\infty$. Indeed $\mathbb{P}(\bigcup_{k \geq \tau_A(\omega)} E_k^c) \leq \sum_k \mathbb{P}(E_k^c)$. But $\exists k_0, \exists C$ such that if $k \geq k_0$, we have

$$\mathbb{P}\left(\frac{U_1 + \cdots + U_k - kE(\mathscr{F})}{k} > \theta - E(\mathscr{F}) + \frac{A - \theta}{k}\right)$$

$$\leq \mathbb{P}\left(\frac{U_1 + \cdots + U_k - kE(\mathscr{F})}{k} > \theta - E(\mathscr{F})\right) \leq C \exp(-kh(\theta - E\mathscr{F}))$$

where h is the Cramer transformed of distribution $\mathscr{F}$. The function h does exist at left since U is left-bounded.

For $k < k_0$,

$$\mathbb{P}(U_1 + \cdots + U_k - k\theta > A - \theta) \leq \varphi_k(A) \quad \text{for some function } \varphi_k.$$

So

$$\mathbb{P}\left(\bigcup_{k \geq \tau_A(\omega)} E_k^c\right) \leq k_0 \max_k \varphi_k(A) + C \sum_{k_0}^{+\infty} e^{-kh(\theta - E\mathscr{F})}.$$

Therefore for every ε, there exists a K such that the second term $< \frac{1}{2}\varepsilon$ if $K < K_0$, and a A_0 such that the first term (with $K_0 = K$) $< \frac{1}{2}\varepsilon$ if $A > A_0$. Hence (4.1.1) holds. Therefore on $E_A^\infty = \bigcap_{k \geq \tau_A} E_k$, the limit distribution of (X_2) given E_A' converges weakly to distribution $\mathscr{F}$, since E_A^∞ has asymptotically probability 1. $\square$

4.2. N-neuron network (divergent case)

We can observe in the simulations that the threshold θ_0 is underestimated by $E(\mathscr{F})/|\mathscr{V}|$. That can be explained by considering the proof of Lemmas 3.1.1 and 3.1.2: We have overestimated the number of discharges that the neighbours of i have by the number that they would have if they discharge according to their own distribution $\mathscr{F}$, i.e. with frequency $1/E(\mathscr{F})$ instead of $1/(E(\mathscr{F}) + \theta|\mathscr{V}|)$. When $\theta > \theta_0$, divergence occurs and maps of active or inhibited neuron do appear: some differences increase to $+\infty$. A stable map of active or inhibited neurons (i.e. of sites i where X_i' remains bounded or sites j where X_j' increases to infinity), divides the network into two subsets E and F such that after a time T the active neurons all belong to E.

Let us give an asymptotical stability criterion for a given map. Suppose that for every $i \in E$, $\mathscr{V}_E(i) = \{i' \in E / i \in \mathscr{V}(i')\}$ has the same cardinal v, and that for every $j \in F$, $\mathscr{W}_E(j) = \{i \in E / j \in \mathscr{V}(i)\}$ is not empty and has a cardinal w_j. Note straightaway that if $w_j = 0$ for some j, j necessarily becomes active after a finite time since it always decreases without being inhibited. From Section 3, we know that if

$$\theta v < E(\mathscr{F}),$$
\hfill (4.2.1)

the process restricted to E is recurrent. If $N_i(t)$ is the counting process at site i (which counts the discharges of neuron i before time t), for every initial distribution

μ on E, and for every site i, it holds

$$\frac{N_i(t)}{t} \xrightarrow[\substack{\text{a.s.} \\ t \to +\infty}]{} \frac{1}{E(\mathcal{F}) + \theta v}.$$

Then for $j \in F$, we define

$$Y_j(t) = \theta \sum_{i \in \mathcal{W}_E(j)} N_i(t)$$

and

$$\frac{Y_i(t)}{t} \xrightarrow[\substack{\text{a.s.} \\ t \to +\infty}]{} \frac{\theta w_j}{E(\mathcal{F}) + \theta v}.$$

So if

$$\forall j \in F, \quad \theta(w_j - v) > E(\mathcal{F}) \tag{4.2.2}$$

there exists t_0 (dependent on μ) such that

$$\mathbb{P}^{\mu}(\forall j \in F, \ Y_j(t) > t) \geq 1 - \varepsilon.$$

By picking at random the initial values X_j^0 on sites $j \in F$ with distribution ν such that $X_j^0 > 2t_0 \ \forall j$ (so one defines a measure $\mu + \nu$ on $E \cup F$), one ensures that no discharge occurs in F before time t_0 (for $t < t_0$, $X_j^t > 2t_0 - t_0 \ \forall j \in F$). So for every $\varepsilon > 0$, there exists an initial distribution $\mu + \nu$ on $E \cup F$, such that with probability $(1 - \varepsilon)$, $(X_i)_{i \in E}^t$ is a recurrent process and $(X_j)_{j \in F}^t \to +\infty$ when $t \to +\infty$.

As every state can be hit after $N + 1$ transitions (Lemma 2.3.3 and Corollary 2.3.4) we get:

Proposition 4.2.1. *Let E, F be a partition of the network with*

$$\mathcal{V}_E(i) = \{i' \in E / i \in \mathcal{V}(i')\} \quad \text{for } i \in E,$$

$$\mathcal{W}_E(j) = \{i \in E / j \in \mathcal{V}(i)\} \quad \text{for } j \in F,$$

and

$$|\mathcal{V}_E(i)| = v, \qquad |\mathcal{W}_E(j)| = w_j > 0 \quad \forall j \in F.$$

So if $\theta v < E(\mathcal{F}) < \theta(w_j - v) \ \forall j \in F$, there is a positive probability that the process $(X_i)_{i \in E}^t$ is recurrent and that

$$(X_j)^t \xrightarrow{\text{a.s.}} +\infty \quad \text{when } t \to +\infty \text{ for each } j \in F. \qquad \square$$

Remark 4.2.1. If E is a set such that $(X_i)_{i \in E}^t$ is recurrent $(\sup_{i \in E} |\mathcal{V}_E(i)| \theta < E(\mathcal{F}))$, then E 'absorbs' every j of F such that $w_j \theta < E(\mathcal{F})$. A stable configuration (E, F) is saturated for the condition

$$\sup_{i \in E} |\mathcal{V}_E(i)| \theta < E(\mathcal{F}) \tag{4.2.3}$$

and verifies

$$\left(\inf_j |\mathcal{W}_E(i)| - \sup_i |\mathcal{V}_E(i)| \right) \theta > E(\mathcal{F}). \tag{4.2.4}$$

(It follows that $2 \sup_i |\mathcal{V}_E(i)| < \inf_j |\mathcal{W}_E(j)|$.) There will be bascule and instability if one of these conditions does not hold. The case $|\mathcal{V}|\theta < E\mathcal{F}$ corresponds to $w_j = v = |\mathcal{V}|$, and to recurrence on E in its entirety.

Remark 4.2.2. The width of a neuron strip in F is limited by the size of the neighbourhood $\mathcal{V}$: Every site in F have to be inhibited by at least one site in E, $w_j > 0 \; \forall j \in F$ and $w_j > v$.

4.3. Examples of observed limit situations

The underlying network is a tore whose periodic local image is represented. Of course, every translation of a limit configuration gives one more.

Let $\times$ denote points in E, and $\circ$ denote points in F. We put

$$\theta_1 = \sup_j \frac{E\mathcal{F}}{w_j - v}, \quad \theta_2 = \frac{E\mathcal{F}}{v}, \quad \text{so } (E, F) \text{ is stable for } \theta_1 < \theta < \theta_2. \qquad (4.3.1)$$

(a) $|\mathcal{V}| = 4$, $\mathcal{V} = \circ \;\; \times \;\; \circ$ (with $\circ$ above and below)

$$(1) \quad \begin{pmatrix} \times & \circ & \times & \circ \\ \circ & \times & \circ & \times \\ \times & \circ & \times & \circ \\ \circ & \times & \circ & \times \end{pmatrix}, \qquad v = 0, \quad w_j = 4, \quad \theta_1 = \tfrac{1}{4}E(\mathcal{F}), \quad \theta_2 = +\infty,$$

$$(2) \quad \begin{pmatrix} \times & \circ & \circ & \times \\ \circ & \times & \circ & \circ \\ \circ & \circ & \times & \circ \\ \times & \circ & \circ & \times \end{pmatrix}, \qquad v = 0, \quad w_j = 2, \quad \theta_1 = \tfrac{1}{2}E(\mathcal{F}), \quad \theta_2 = +\infty.$$

Note. If the studied tore is $\mathbb{T} = \mathbb{Z}/2m \times \mathbb{Z}/2m'$, the particular shape of neighbourhoods leads to a check pattern: $\mathbb{T} = E + F$ with $\mathcal{V}(i) \subset E$ for $i \in F$ and vice versa. But however, we get a limit situation (case (2)) which is not check-shaped. Notice that for $\theta > \tfrac{1}{2}E(\mathcal{F})$, both kinds are final states.

(b) $|\mathcal{V}| = 6$, $\mathcal{V} = \circ \;\; \circ \;\; \times \;\; \circ \;\; \circ$ (with $\circ$ above and below the $\times$)

$$(1) \quad \begin{pmatrix} \circ & \circ & \times & \times & \circ & \circ \\ \times & \times & \circ & \circ & \times & \times \\ \circ & \circ & \times & \times & \circ & \circ \\ \times & \times & \circ & \circ & \times & \times \end{pmatrix}, \qquad v = 1, \quad w_j = 5, \quad \theta_1 = \tfrac{1}{4}E(\mathcal{F}), \quad \theta_2 = E(\mathcal{F}),$$

$$(2) \quad \begin{pmatrix} \times & \circ & \circ & \times & \circ & \circ & \times \\ \circ & \times & \circ & \circ & \times & \circ & \circ \\ \circ & \circ & \times & \circ & \circ & \times & \circ \\ \times & \circ & \circ & \times & \circ & \circ & \times \end{pmatrix}, \quad v = 0, \quad w_j = 2, \quad \theta_1 = \tfrac{1}{2}E(\mathcal{F}), \quad \theta_2 = +\infty,$$

$$(3) \quad \begin{pmatrix} \times & \circ & \circ & \circ & \times & \circ & \circ & \circ \\ \circ & \times & \circ & \circ & \circ & \times & \circ & \circ \\ \circ & \circ & \times & \circ & \circ & \circ & \times & \circ \\ \circ & \circ & \circ & \times & \circ & \circ & \circ & \times \end{pmatrix}, \quad \begin{array}{l} v = 0, \quad w_j = 1 \text{ or } 2, \\[6pt] \theta_1 = E(\mathcal{F}), \quad \theta_2 = +\infty. \end{array}$$

$$(4) \quad \begin{pmatrix} \times & \circ & \circ & \circ & \circ & \times & \circ & \circ \\ \circ & \times & \circ & \circ & \circ & \circ & \times & \circ \\ \circ & \circ & \times & \circ & \circ & \circ & \circ & \times \\ \circ & \circ & \circ & \times & \circ & \circ & \circ & \circ \end{pmatrix}, \quad \begin{array}{l} v = 0, \quad w_j = 1 \text{ or } 2, \\[6pt] \theta_1 = E(\mathcal{F}), \quad \theta_2 = +\infty. \end{array}$$

Moreover, see in Figures 5, 6 that they are actually limit situations when the process X^t is initialized at random. It would remain to determine how the choice between the stable maps is made, when various stable maps do exist for one value of θ and to study the bifurcation phenomenon.
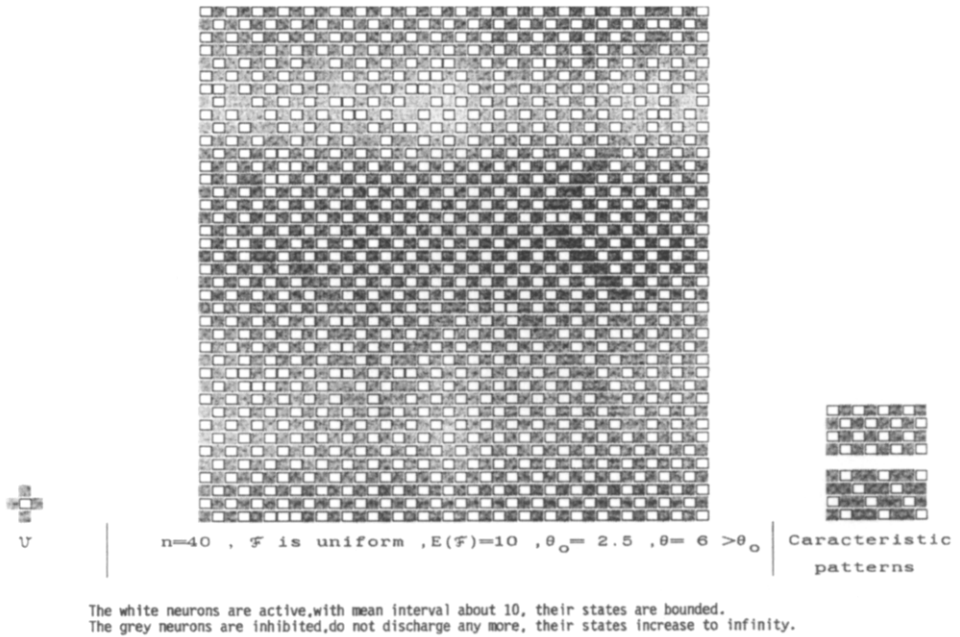

n=40 , $\mathcal{F}$ is uniform ,E($\mathcal{F}$)=10 ,θ_0= 2.5 ,θ= 6 >θ_0 Caracteristic
                                                                                                patterns

The white neurons are active,with mean interval about 10, their states are bounded.
The grey neurons are inhibited,do not discharge any more, their states increase to infinity.

Fig. 5. 4 neighbours, system state after 2000 steps.

U | n=40 , 𝔉 is uniform ,E(𝔉)=10 ,θ₀= 1.66,θ= 6 >θ₀ | Caracteristic patterns

The white neurons are active,with mean interval about 10, their states are bounded.
The grey neurons are inhibited,do not discharge any more, their states increase to infinity.
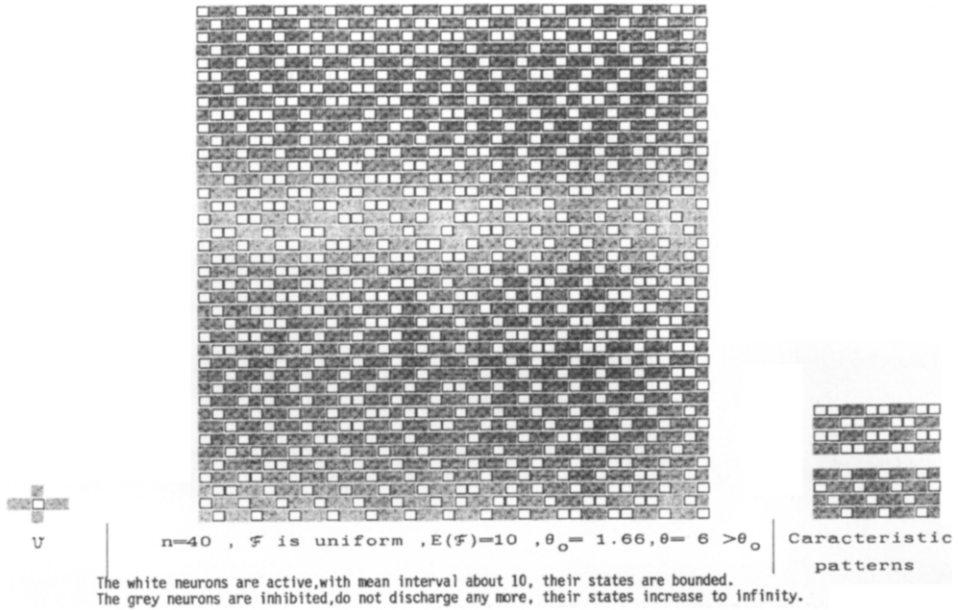
Fig. 6. 6 neighbours, system state after 2000 steps.

5. Stimulated network

In this section, we take external stimuli into account, and suppose that some neurons are excited. (In the cerebellar cortex, climbing fibres and mossy fibres stimulate the Purkinjie cells.)

5.1. The model

So let us consider that the activity of neuron i, if it was isolated from the others, is represented by a renewal process with an interval distribution $\mathscr{F}_i$ which depends on i. Neuron i is more excited than neuron j if and only if $E(\mathscr{F}_i) < E(\mathscr{F}_j)$. We suppose that distributions $\mathscr{F}_i$ have positive density on $\mathbb{R}^+$ and finite mean values and variances, as in Section 2.

Network state X^t is defined in the same way as in Section 2.1 (by substituting $\mathscr{F}_i$ for $\mathscr{F}$). Markov chain X^{T_l}, that is sampled at discharge instants T_l, and Markov process X^t are irreducible and aperiodic. Their behaviour depends on the inhibition parameter θ.

5.2. Two-neuron system

Denote by U with distribution $\mathscr{F}_1$ (resp. V with distribution $\mathscr{F}_2$) the r.v. which is used to reinitialize the state of neuron 1 (resp. 2) just after each discharge. Suppose $E(\mathscr{F}_1) < E(\mathscr{F}_2)$, that is neuron 1 is excited. If

$$D^l = (X_2 - X_1)^l = (X_2 - X_1)^{T^l},$$

we have

$$D^{l+1} = D^l + \mathbf{1}_{D^l>0}(\theta - U_{l+1}) + \mathbf{1}_{D^l<0}(V_{l+1} - \theta).$$

(U_{l+1} and V_{l+1} are independent of X^{T_l}, and have distribution $\mathscr{F}_1$ and $\mathscr{F}_2$ respectively.)
So:
 (a) If $\theta < E(\mathscr{F}_1)$, mean jump of D^l and D^l have opposite signs.
 (b) If $E(\mathscr{F}_1) < \theta < E(\mathscr{F}_2)$, mean jump of D^l is positive.
 (c) If $E(\mathscr{F}_2) < \theta$, mean jump of D^l and D^l have the same sign.
 Hence:

Proposition 5.2.1 (Two-neuron case). *Suppose* $E(\mathscr{F}_1) < E(\mathscr{F}_2)$. *Let* $\theta_0 = \min(E\mathscr{F}_1, E\mathscr{F}_2)$, $\theta_1 = \max(E\mathscr{F}_1, E\mathscr{F}_2)$.
 (i) *When* $\theta < \theta_0$, *both processes* X^{T_l} *and* X^l *are ergodic.*
 (ii) *When* $\theta_0 < \theta < \theta_1$, $X_2^t \xrightarrow[t \to +\infty]{\text{a.s.}} +\infty$ *and* X_1^t *converges in law to a renewal process with distribution* $\mathscr{F}_1$.
 (iii) *When* $\theta > \theta_1$, *both processes* X^{T_l} *and* X^l *diverge. On almost all trajectories, one of the components tends towards* $+\infty$ *a.s., the other converges in law to a renewal process with distribution* $\mathscr{F}_1$ *or* $\mathscr{F}_2$ *(according to the site).*

Proof. (i) and (iii) are proved in Sections 3 and 4.
 For (ii), observe that D^l is a random walk with increment $V - \theta$ and drift $\theta_1 - \theta > 0$ as long as $D^l < 0$, and with increment $\theta - U$ and drift $\theta - \theta_0 > 0$ when $D^l > 0$. □

As in Section 3, the interspike interval for neuron (2) can be written

$$\tau = V_1 + K\theta$$

where K is an integer-valued variable defined by

$$\mathbb{P}(K > k) = \mathbb{P}(V_1 > x_0 + \theta, \ U_1 < V_1 - x_0, \ U_1 + U_2 < V_1 - x_0 + \theta, \ldots,$$

$$U_1 + \cdots + U_k < V_1 - x_0 + (k-1)\theta)$$

where $(x_0, 0)$ is the state of the network just before neuron (2) does discharge, $V_1 \sim \mathscr{F}_2$ and $U_1, \ldots, U_k$ are independent with distribution $\mathscr{F}_1$.
 If t_0, t_1 are consecutive instants of discharge for neuron (2), and X_0, X_1 are the states of neuron (1) at these instants, we have

$$X_1 = U_1 + U_2 + \cdots + U_k - V_1 + X_0 - (K-1)\theta$$

(notations of Section 3.3) where V_1 is the random variable drawn in neuron (2) at instant t_0, $U_1, \ldots, U_k$ are the random variables drawn in neuron (1) between t_0 and t_1.
 In stationary behaviour, $X_1 \sim X_0$ and, as K is a stopping time, we get

$$(EK)(EU - \theta) = EV - \theta.$$

So

$$EK = \frac{EV - \theta}{EU - \theta}.$$

It follows that:

Proposition 5.2.2 (Two-neuron case). *If* $\theta < \theta_0 = \min(E\mathcal{F}_1, E\mathcal{F}_2)$, *the mean interspike intervals are*

$$E\mathcal{F}_2 + \frac{E\mathcal{F}_2 - \theta}{E\mathcal{F}_1 - \theta}\, \theta \quad \text{for neuron (2)},$$

$$E\mathcal{F}_1 + \frac{E\mathcal{F}_1 - \theta}{E\mathcal{F}_2 - \theta}\, \theta \quad \text{for neuron (1)}. \qquad \square$$

5.3. N-neuron system

We can generalize the results of Sections 3 and 4.

Proposition 5.3.1 (Ergodicity). *If* $\theta < \theta_0 = \min_i(E\mathcal{F}_i/|\mathcal{V}|)$, *processes* X^{T_i} *and* X^t *are ergodic.* $\quad \square$

In this case, all neurons are active, and discharge more slowly than if they were isolated. We observe numerically that the mean interspike interval for neuron i is $E\mathcal{F}_i + |\mathcal{V}|\theta$ if i lies inside or outside of the excited zone and

$$E\mathcal{F}_i + \frac{E\mathcal{F}_{i'} - \theta}{E\mathcal{F}_i - \theta}\, \theta$$

if i is at the boundary between the excited and non-excited zones.

Proposition 5.3.2 (Divergence). *When* $\theta > \theta_1 = \max_i(E\mathcal{F}_i/|\mathcal{V}|)$, *process* X^t *diverges.*
$\square$

Its behaviour is the same as in the case without stimulation: inhibited and active neuron strips do appear, which break the lateral inhibition around each neuron. Of course, the mean interspike interval for each active neuron i is approximately equal to its own interval $E\mathcal{F}_i$.

At last, when $\theta_0 < \theta < \theta_1$, some neurons are inhibited, others no, and various maps can be obtained.

Let us examine some cases:

(1) *Only one neuron i_0 is excited*: $\mathcal{F}_i = \mathcal{F}$ for all $i \neq i_0$, and $E\mathcal{F}_{i_0} < E\mathcal{F}$. Then if

$$E\mathcal{F}_{i_0}/|\mathcal{V}| < \theta < E\mathcal{F}/|\mathcal{V}|,$$

only the neighbours of i_0 are inhibited. It is the lateral inhibition.

(2) *One zone D (row or plate) is excited*:

$$E\mathscr{F}_i = \begin{cases} E\mathscr{F}_1 & \text{if } i \in D, \\ E\mathscr{F}_2 & \text{if } i \notin D, \end{cases}$$

with $E\mathscr{F}_1 < E\mathscr{F}_2$. Then if

$$E\mathscr{F}_1/|\mathcal{V}| < \theta < E\mathscr{F}_2/|\mathcal{V}|,$$

there is

convergence outside of D, and divergence inside of D

and at the boundary of D, there are quite inhibited neurons.

6. Conclusion

For delay θ (which models the inhibition due to neighbours' discharge), whose value is in conformity with the neurobiological reality (θ is about 3 or 4 ms, for $E(\mathscr{F})$ approximately 100 ms), one observes a stationary behaviour, according to the theory. We estimated the threshold θ_0 under which the stationary behaviour occurs, even if it is underestimated.

When the inhibition is strong ($\theta > \theta_0$), the network diverges. The interesting phenomenon is the apparition of alternate strips made of active or inhibited neurons. The shape and orientation of these strips depend on the neighbourhood's shape, on its width and on the value θ. When θ increases, a bifurcation phenomenon does occur: new stable situations appear, without disappearance of the previous ones (see especially the case with 6 neighbours).

We could have thought that a lateral inhibition area would take place around each active neuron, to cut them off from one another. On the contrary, we observe that these isolated areas burst by increasing the inhibition parameter, and that moiré responses do appear, which are quite similar to those that we can see in the cortex tissues when they are stimulated: preferential orientation maps, occular dominance in the visual cortex for example. The results can be compared to those of Amari (1977), who gives a taxonomy of neural fields connected by lateral inhibition using deterministic field equations.

Our model brings to the fore in a very clear way the fact that these differentiations are caused by *the inhibition*.

References

S. Amari, Dynamic of pattern formation in lateral-inhibition type neural fields, Biol. Cybern. 27 (1977) 77–87.

H. Axelrad, C. Bernard, M. Cottrell and B. Giraud, The use of an artificial neural network to analyse the information transfer properties of a simplified model of the cerebellar cortex, in: Proc. of the IEEE First Ann. Internat. Conf. on Neural Networks, June 1987 (1987).

H. Axelrad, C. Bernard, B. Giraud and M.E. Marc, Cortex cerebelleux: Simulation de l'inhibition par collatérales récurrentes à l'aide d'un ensemble d'automates neuronaux. I. Modèles et résultats globaux, C. R. Acad. Sci. Sér. III 301(1) (1985).

G.G. Blasdel and G. Salama, Voltage-sensitive dyes reveal modular organization in monkey striate cortex, Nature 321 (1986) 579.

G. Chauvet, Habitation rules for a theory of the cerebellar cortex, Biol. Cybern. 55 (1986) 201–209.

M. Cottrell, Stability and attractivity in associative memory networks, Biol. Cybern. 58 (1988a) 129–139.

M. Cottrell, Analysis of a simplified model of the cerebellar cortex, in: G. Dreyfus, ed., Proc. of n'Euro88 Conf. (1988b).

M. Cottrell and J.C. Fort, A stochastic model of retinotopy: A self organizing process, Biol. Cybern. 53 (1986) 405–411.

M. Cottrell and J.C. Fort, Etude d'un processus d'auto-organisation, Ann. Inst. H. Poincaré 23(1) (1987) 1–20.

D.R. Cox and P.A.W. Lewis, Statistical Analysis of Series of Events (Methuen, London, 1968).

F. Crepel, N. Delhaye-Bouchaud, J.L. Dupont and C. Sotelo, Dentritic and axonic fields of Purkinje cells in developing and X-irradiated rat cerebellum. A comparative study using intracellular staining with horse radish peroxidase, Neuro-Sci. 5 (1980) 333–347.

J.C. Eccles, M. Ito and J. Szentagothai, The cerebellum as a Neuronal Machine (Springer, Berlin, 1967).

B. Giraud, C. Bernard and H. Axelrad, Cortex cerebelleux: simulation de l'inhibition par collatérales récurrentes à l'aide d'un ensemble d'automates neuronaux. II. Modulation temporelle et contenu informationnel, C. R. Acad. Sci. Paris Sér. III 301(11) (1985).

M. Ito, The Cerebellum and Neural Control (Raven Press, New York, 1984).

T. Kohonen, Self-organization and Associative Memory (Springer, Berlin, 1984).

J.P. Landolt and M.J. Correia, Neuromathematical concepts of point process theory, IEEE Trans. BME25 (1978) 1–12.

W.S. MacCulloch and W. Pitts, A logical calculus of ideas immanent in nervous activity, Bull. Math. Biophys. 5 (1943) 115–133.

D. Marr, A theory of cerebellar cortex, J. Physiol. 202 (1969).

J. Neveu, Martingales à Temps Discret (Masson, Paris, 1972).

E. Nummelin, General Irreducible Markov Chains and Non-negative Operators (Cambridge Univ. Press, Cambridge, 1984).

S.L. Palay and V. Chan-Palay, Cerebellar Cortex (Springer, Berlin, 1974).

D. Revuz, Markov Chains (North-Holland, Amsterdam, 1975).

T.E. Rumelhart and J.L. McClelland, Parallel Distributing Processing, Tome 1 et 2 (MIT Press, Cambridge, MA, 1986).

G. Sampath and S.K. Srinivasan, Stochastic Models for Spike Trains of Single Neurons (Springer, Berlin, 1977).

J. Von Neumann, Probabilistic logics and the synthesis of reliable organisms from unreliable components, in: C.E. Shannon and J. McCarthy, eds., Automata Studies (Princeton Univ. Press, Princeton, NJ, 1956) pp. 43–98.