INT. J. SYSTEMS SCI., 1983, VOL. 14, NO. 10, 1191–1216

Time evolution of a system of two valued interacting elements: a microscopic interpretation of birth and death equations

BERTRAND ROEHNER†

We consider here one of the simplest possible systems with N interacting particles. It has the following features: (i) the state variable of each particle takes the values $\sigma_i = \mp 1$; (ii) the interaction is chosen in such a way to preserve the symmetry of the distribution function $p(\sigma_1, \sigma_2, \ldots, \sigma_N; t)$ with respect to the σ_i; and (iii) the evolution of the system is defined in a stochastic way by the transition probabilities of each particle as depending on the state of all other particles. The master equation of this Markov process is shown to be the equation of a general birth and death process in one dimension. More precisely, the birth and death process is: linear if the particles are independent; quadratic if there is a binary interaction; or cubic if there is a third-order interaction. We develop the reduced distribution equations hierarchy (which is the analogue of the BBGKY hierarchy) and we study under what conditions this hierarchy closes. Then we show that for specific systems there is a conserved quantity (in the mean) and we discuss for what kind of interaction there is respectively an H-theorem and a postulate of equal *a priori* probabilities at equilibrium. It appears in particular that this postulate should not be true in the strong form in which it is usually stated.

1. Introduction

Understanding the collective behaviour and time dependency of systems involving many interacting elements is the main task of non-equilibrium statistical mechanics. This question is also crucial in other fields, as indicated in the following.

There are always two approaches open to us: we may build solvable models from which, despite of their oversimplification of reality, something can nevertheless be learned (at least on a qualitative level); or we may try to approximate real systems as closely as possible and resort to approximate or numerical solution. As far as interacting systems are concerned, it is well known that the mathematical difficulties become very serious even for the simplest models, unfortunately making the gap between solvable models and realistic ones very large.

In this article, we shall try the first approach. Actually we are not considering a specific model, rather a class of models. The distinctive features of these models are: the elements of the system are two valued spins, although the generalization to three and higher valued spins will be indicated; and the spins will be considered as playing identical roles in the system. Thus, it is representative of a gas rather than a solid state lattice.

The equation giving the most complete description, that is the Liouville equation of the system, turns out to be the Kolmogorov equation of a one dimensional birth and death process.

Received 21 December 1982.

† Laboratoire de Physique Théorique et Hautes Energies, Université Paris VII, Tour 24, 5 1étage, 2 place Jussieu, 75251 Paris, Cedex 05, France.

A very stimulating investigation of the time evolution of interacting spins was presented by Glauber (1963). However the emphasis was rather on lattice spins (in relation with the Ising model) involving interactions between nearest neighbours. The Liouville equation of this N spins system was thus an N-dimensional birth and death process.

On the other hand, Kac (1967) studied, from the point of view of non-equilibrium statistical mechanics, a specific model of spin-gas : the so called MacKean model. This particular case will be considered later.

For the present model, the Liouville equation will more specifically be the Kolmogorov equation of a linear birth and death process when there is no interaction between the elements of the system, a quadratic birth and death process (i.e. a birth and death process with quadratic transition rates λ_n, μ_n) when there is a binary interaction between the elements, or a p-degree birth and death process when p-order interactions are taken into account.

Of course, these birth and death equations are not equally easy to solve. The linear birth and death process has a well known and easy solution (Bailey 1964, p. 90) : since it corresponds to a system of independent particles, this is not surprising. For quadratic processes, an exact solution is known only in specific cases (see § 2). Finally, no solvable process of degree higher than two is known. Since the case of binary interactions is by far the most important, it points out the prominent part played by quadratic birth and death processes.

Besides this microscopic interpretation of birth and death processes, the relative simplicity of the Liouville equation will enable us to make observations about the basic postulates and ideas of statistical mechanics.

Before considering the evolution of a stochastic system, it is usually a good policy to first look at the corresponding deterministic system. However, this is not easy here since there is no convenient mathematical description of a deterministic system oscillating between two states.

The paper has the following structure. In § 2 we develop the Liouville equation of our system. We then solve completely a simple example before coming to the general connection with birth and death processes. We indicate also how our reasoning could be extended to spins taking more than two values. Then we write the equations for the mean values and the reduced distribution functions hierarchy. In § 3, we discuss the following points : the existence of conserved quantities ; the possibility of an H-theorem ; and the validity of the postulate of equal *a priori* probabilities at equilibrium. Before proceeding, let us classify interacting systems and take some examples from physics, neurophysiology and economics.

1.1. *Classification of interacting systems*

Let $\mathbf{x}_i$ denote the phase space coordinates of the ith element of a system : $1 \leqslant i \leqslant N$. Let $w_i(\mathbf{x}_1, \ldots, \mathbf{x}'_i, \ldots, \mathbf{x}_N/\mathbf{x}_1, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_N)$ denote the transition rate from the state $(\mathbf{x}_1, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_N)$ to the state $(\mathbf{x}_1, \ldots, \mathbf{x}'_i, \ldots, \mathbf{x}_N)$, $\mathbf{x}_i \neq \mathbf{x}'_i$. We shall say that an interaction is a global one element interaction if

$$w_i(\mathbf{x}_1, \ldots, \mathbf{x}'_i, \ldots, \mathbf{x}_N/\mathbf{x}_1, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_N) = \sum_{i=1}^{N} v(\mathbf{x}_i, \mathbf{x}'_i)$$

a global binary interaction if

$$w_i(\mathbf{x}_1, \ldots, \mathbf{x}'_i, \ldots, \mathbf{x}_N/\mathbf{x}_1, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_N) = \sum_{\substack{i,j \\ j \neq i}}^{N} v(\mathbf{x}_i, \mathbf{x}'_i\,;\ \mathbf{x}_j)$$

a global three order interaction if

$$w_i(\mathbf{x}_1, \ldots, \mathbf{x}'_i, \ldots, \mathbf{x}_N/\mathbf{x}_1, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_N) = \sum_{\substack{i,j,k \\ i \neq j,k \\ j \neq k}} v(\mathbf{x}_i, \mathbf{x}'_i\,;\ \mathbf{x}_j, \mathbf{x}_k)$$

We shall say that an interaction is a local binary interaction if

$$w_i(\mathbf{x}_1, \ldots, \mathbf{x}'_i, \ldots, \mathbf{x}_N/\mathbf{x}_1, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_N) = \sum_{\substack{j \in v(i) \\ j \neq i}} v(\mathbf{x}_i, \mathbf{x}'_i\,;\ \mathbf{x}_j)$$

where $v(i)$ denotes the indices set of the elements located in a specific neighbourhood of $\mathbf{x}_i$. A similar definition holds of course for local one or three bodies interactions.

1.2. *Examples*

Which of those interactions are most common in the three fields mentioned above ?

In solid state physics, the particles of the system are subject to only small vibrations about their fixed equilibrium positions on the lattice. Moreover, the interaction is mainly with a few nearest neighbours and takes the form of a binary interaction. The interaction is thus a local binary interaction. For a fluid (gas or liquid) however, due to the mobility of the particles and to possible long range interaction (especially for electrically charged particles), every particle may interact with every other. Thus we have a global binary interaction.

In the human brain considered as a system, the elements are the neurones. Crudely speaking the neurones have two states : excited or not excited. They do not have any mobility but each of them is in connection with hundreds of thousands (let p denote this number) of other neurones. We have thus a situation which is intermediate between the global and local interaction. For the interaction to be global, at least in some functional part of the brain, the number of connections of every neurone should be of the order of magnitude of the total number of neurones, that is close to one billion. This is not the case.

On the other hand, a local interaction with so many neighbours is difficult to formalize. The global interaction could possibly be locally a reasonable approximation for distances less than $p^{1/3}v \sim 100a \sim 1$ mm (where a is the dimension of a neurone). For larger distances, there are macroscopic functional patterns and paths (the olfactive or visual nerve for example).

In the world economic network, the elements are the enterprises. For a span of time of the order of several years, they can be considered as geographically fixed. The number of enterprises to which one of them is connected depends of course on its size, but the order of magnitude of this number is a fraction of one hundred†. The total number and the number

† A clear indication of those connections is given for the different sectors of economic activity by input–output tables.

of connecting elements is about four of five orders of magnitude smaller than in the previous case, but the same crude considerations seem to apply in respect to the kind of interaction.

In this paper, we shall be concerned exclusively with global interactions. From the previous discussion it appears that this kind of interaction occurs for dense gases and liquids and also locally in the brain or in the economic world.

2. Microscopic interpretation of birth and death processes

2.1. *Basic equations*

2.1.1. *Notations*

Consider N elements taking the values : $\sigma_i = \mp 1$. The whole system has 2^N possible states and each state is completely described by the vector

$$\boldsymbol{\sigma} = (\boldsymbol{\sigma}_1, \ldots, \boldsymbol{\sigma}_N)$$

We make the following assumptions for the transitions of this stochastic system between time t and $t+\Delta t$. First, one element is drawn at random from the set $\sigma_1, \ldots, \sigma_N$: let i be its index. This element will experience a transition from its present state σ_i to state $-\sigma_i$ with the probability

$$N w_i(\sigma_1, \ldots, \sigma_i, \ldots, \sigma_N) \cdot \Delta t$$

where w_i is a positive function of its arguments.

The following notation will be useful

$$\boldsymbol{\sigma}_i = (\sigma_1, \ldots, -\sigma_i, \ldots, \sigma_N)$$

The probability of a transition from a given state $\boldsymbol{\sigma}$ to a different one (i.e. to any of the $\boldsymbol{\sigma}_i$) between time t and $t+\Delta t$ is

$$\Delta t \cdot \sum_{i=1}^{N} w_i(\boldsymbol{\sigma})$$

We denote the probability for the system to be in state $\boldsymbol{\sigma}$ at time t as $p(\boldsymbol{\sigma}\,;\, t)$.

2.1.2. *The Kolmogorov equation*

To get a differential equation for $p(\boldsymbol{\sigma}\,;\, t)$ we write the probability conservation relation between time t and $t+\Delta t$

$$p(\boldsymbol{\sigma}\,;\, t+\Delta t) = p(\boldsymbol{\sigma}\,;\, t)\left(1 - \Delta t \sum_{i=1}^{N} w_i(\boldsymbol{\sigma})\right) + \Delta t \cdot \sum_{i=1}^{N} w_i(\boldsymbol{\sigma}_i) p(\boldsymbol{\sigma}_i\,;\, t)$$

As Δt goes to zero, the Kolmogorov equation of this discrete Markov process is obtained, i.e.

$$\left.\begin{aligned} \dot{p}(\boldsymbol{\sigma}\,;\, t) &= -p(\boldsymbol{\sigma}\,;\, t) \sum_{i=1}^{N} w_i(\boldsymbol{\sigma}) + \sum_{i=1}^{N} w_i(\boldsymbol{\sigma}_i) p(\boldsymbol{\sigma}_i\,;\, t) \\ \dot{p} &= \frac{dp}{dt}\,; \quad p(\boldsymbol{\sigma}\,;\, t=0) = p_0(\boldsymbol{\sigma}) \end{aligned}\right\} \qquad (1)$$

This equation is also referred to as the master equation of the process (Van Kampen 1962). We may say that it is the Liouville equation of the system for the following reasons : the Liouville equation is usually derived from the evolution equation

$$\dot{\mathbf{x}}=\mathbf{f}(\mathbf{x}\ ;\ t)$$

taking into account the property that in classical mechanics

$$\operatorname{div}\mathbf{f}=0$$

However, it simply expresses the probability conservation between time t and $t+\Delta t$ and it can be written without any additional condition on $\mathbf{f}$, in which case it reads

$$\frac{\partial F}{\partial t}+\sum_{i=1}^{N}\frac{\partial F}{\partial x_i}\dot{x}_i+\operatorname{div}\mathbf{f}=0 \tag{2}$$

$F(\mathbf{x}\ ;\ t)\,d\mathbf{x}$ is the probability of finding the system in the neighbourhood $d\mathbf{x}$ of $\mathbf{x}$ at time t. Besides the fact that the state variable is continuous in one case and discrete in the other, the only difference between eqns. (1) and (2) lies in the origin of the stochastic behaviour : for eqn. (2), the evolution is deterministic but the initial condition is a probabilistic one, while for eqn. (1), the evolution itself is stochastic (the initial condition may be either deterministic or probabilistic).

It can be argued that eqn. (2) describes a reversible process, whereas the process described by (2) is usually irreversible (see § 3.2). This difference is however only formal. It is well known (Rice 1965, p. 173) that the coarse grained Liouville equation (which is the only one having an experimental meaning) derived from the fine grained equation (2), is irreversible too†.

The reason for the previous discussion is, of course, not only one of nomenclature : it should emphasize that starting from eqn. (1), we will follow the usual steps of non-equilibrium classical mechanics comprising, for example, the BBGKY hierarchy, the Boltzmann equation, etc.

2.2. *Particular cases*

2.2.1. *Independent particles*

As a simple example we assume first the $w_i(\boldsymbol{\sigma})$ to be independent of $\boldsymbol{\sigma}$. Supposing the initial condition is such that $p(\boldsymbol{\sigma}\ ;\ t=0)$ is a symmetric function of the variables $\sigma_1, \ldots, \sigma_N$, then $p(\boldsymbol{\sigma}\ ;\ t)$ will be symmetric too. To prove this property it is sufficient to show that if $p(\boldsymbol{\sigma}\ ;\ t)$ is symmetric, then $p(\boldsymbol{\sigma}\ ;\ t+\Delta t)$ is symmetric also. Now

$$p(\boldsymbol{\sigma}\ ;\ t+\Delta t)=p(\boldsymbol{\sigma}\ ;\ t)(1-N\cdot\Delta t\cdot w)+w\Delta t\sum_{i=1}^{N}p(\boldsymbol{\sigma}_i\ ;\ t)$$

The last term is obviously symmetric since the interchange of σ_i and σ_j will only interchange $p(\boldsymbol{\sigma}_i\ ;\ t)$ and $p(\boldsymbol{\sigma}_j\ ;\ t)$.

---

† The well known fact (Van Kampen 1962, p. 117) that the trajectories of the system in phase space are very ragged is another way to express the similarity between random and deterministic systems of a large number of particles.

Now a function depending symmetrically on two valued variables is a function solely of their sum (Appendix 1)

$$p(\boldsymbol{\sigma}\,;\ t)=q(\sigma_1+\ldots+\sigma_N\,;\ t)=p_n(t)\Big/\binom{N}{n}$$

n denotes the number of ones in the sequence $\sigma_1, \ldots, \sigma_N$ and $p_n(t)$ is the probability that n of the σ_i are equal to one at time t.

Due to the symmetry of the function $p(\boldsymbol{\sigma}\,;\ t)$, we may assume that its arguments are always ordered in such a way that the numbers one precede the numbers minus-one. Thus, the Kolmogorov equation becomes

$$\dot{p}_n\Big/\binom{N}{n}=-Nw\left[p_n\Big/\binom{N}{n}\right]+wn\left[p_{n-1}\Big/\binom{N}{n-1}\right]$$
$$+w(N-n)\left[p_{n+1}\Big/\binom{N}{n+1}\right]$$

resulting in the following linear (and finite) birth and death process

$$\dot{p}_n=-Nwp_n+w[N-(n-1)]p_{n-1}+w(n+1)p_{n+1} \tag{3}$$

The transition rates are

$$\lambda_n=w(N-n)$$

$$\mu_n=wn$$

Remarks

(1) The case $w_j(\boldsymbol{\sigma})=w$ is not the most general leading to a linear birth and death process (see § 2.3).

(2) Although this case is almost trivial, it serves to indicate the steps which will be the same in more complicated situations.

2.2.2. *An infinite linear process*

What do we have to modify in our assumptions to get an infinite linear process instead of a finite one ?

If we let $N\to\infty$ in eqn. (3) we observe that the birth term diverges. This is as expected. Indeed if we start with a (finite) number n of ones, letting $N\to\infty$ introduces an infinite number of minus-ones ; this makes the transitions $-1\to1$ far more frequent than the opposite transitions. To render the limiting process $N\to\infty$ possible, we have to weight differently both transitions. Suppose that

$$w_i(\sigma_1, \ldots, \sigma_{i-1}, -1, \sigma_{i+1}, \ldots, \sigma_N)=\lambda\,\frac{n}{N-n}$$

$$w_i(\sigma_1, \ldots, \sigma_{i-1}, 1, \sigma_{i+1}, \ldots, \sigma_N)=\mu$$

Since the new transition rates depend only on the sum

$$\sigma_1 + \ldots + \sigma_N$$

it is clear that we again have the symmetry conservation property. By the same reasoning as before, we obtain the following equation for $p_n(t)$

$$\dot{p}_n = -(\lambda+\mu)np_n + \lambda(n-1)p_{n-1} + \mu(n+1)p_{n+1}$$

Since this equation no longer involves N, the limiting process can be performed without trouble. Of course, the elements of this system are no longer independent : their transitions are sensitive to the compounding of the whole population.

In the sequel, we will be concerned with finite birth and death processes which arise in a more natural way. The previous reasoning could however be repeated leading in each case to a microscopic interpretation of the corresponding infinite birth and death process.

2.2.3. *Generalization to elements taking more than two values*

Assume, for definiteness, that σ_i may take the values : $-1, 0, 1$. The expression of the transition rate must now indicate both the initial value and the final one. We denote by

$$w_i(\sigma, \sigma' ; \boldsymbol{\sigma})\Delta t$$

the probability that σ_i will change from σ to σ' between t and $t+\Delta t$. The Kolmogorov equation will thus read

$$\dot{p}(\boldsymbol{\sigma} ; t) = -\left[\sum_{i=1}^{N} \sum_{\sigma \neq \sigma_i} w_i(\sigma_i, \sigma ; \boldsymbol{\sigma})\right] p(\boldsymbol{\sigma}, t) + \sum_{i=1}^{N} \sum_{\sigma \neq \sigma_i} w_i(\sigma, \sigma_i ; \boldsymbol{\sigma}_i(\sigma)) p(\boldsymbol{\sigma}_i(\sigma) ; t) \quad (4)$$

where

$$\boldsymbol{\sigma}_i(\sigma) = (\sigma_1, \ldots, \sigma_{i-1}, \sigma, \sigma_{i+1}, \ldots, \sigma_N)$$

This time the symmetric function $\dot{p}(\boldsymbol{\sigma} ; t)$ will depend on (Appendix 1)

$$s = \sigma_1 + \ldots + \sigma_N$$

$$r = \sigma_1^2 + \ldots + \sigma_N^2$$

Let n_1 denote the number of ones and n_0 the number of zero. Hence

$$p(\boldsymbol{\sigma} ; t) = q(s, r ; t) = p(n_0, n_1, n_2 ; t) \Big/ \binom{N}{n_0 n_1 n_2} \quad (5)$$

The variable $n_2 = N - n_0 - n_1$ is superfluous but it makes the equations more symmetric.

Replacing $p(\boldsymbol{\sigma}\,;\,t)$ in eqn. (4) by its expression (5) gives

$$
\begin{aligned}
\dot{p}(n_0, n_1, n_2\,;\,t) = & -p(n_0, n_1, n_2\,;\,t) \sum_{i=1}^{N} \sum_{\sigma \neq \sigma_i} w_i(\sigma_i, \sigma\,;\,\boldsymbol{\sigma}) \\
& + \frac{n_1+1}{n_0} p(n_0-1, n_1+1, n_2) \sum_{i=1}^{n_0} w_i(1, 0\,;\,\boldsymbol{\sigma}_i(1)) \\
& + \frac{n_2+1}{n_0} p(n_0-1, n_1, n_2+1) \sum_{i=1}^{n_0} w_i(-1, 0\,;\,\boldsymbol{\sigma}_i(-1)) \\
& + \frac{n_0+1}{n_1} p(n_0+1, n_1-1, n_2) \sum_{i=n_0+1}^{n_0+n_1} w_i(0, 1\,;\,\boldsymbol{\sigma}_i(0)) \\
& + \frac{n_2+1}{n_1} p(n_0, n_1-1, n_2+1) \sum_{i=n_0+1}^{n_0+n_1} w_i(-1, 1\,;\,\boldsymbol{\sigma}_i(-1)) \\
& + \frac{n_0+1}{n_2} p(n_0+1, n_1, n_2-1) \sum_{i=n_0+n_1+1}^{N} w_i(0, -1\,;\,\boldsymbol{\sigma}_i(0)) \\
& + \frac{n_1+1}{n_2} p(n_0, n_1+1, n_2-1) \sum_{i=n_0+n_1+1}^{N} w_i(1, -1\,;\,\boldsymbol{\sigma}_i(1))
\end{aligned}
$$

As an application of this lengthy formula, we specialize to the case of independent particles

$$w_i(\sigma, \sigma'\,;\,\boldsymbol{\sigma}) = w$$

Introducing the generating function

$$G(x_0, x_1, x_2\,;\,t) = \sum_{\substack{n_0, n_1, n_2 \\ n_0+n_1+n_2=N}} x_0^{n_0} x_1^{n_1} x_2^{n_2} p(n_0, n_1, n_2\,;\,t)$$

the previous system is seen to be equivalent to the partial differential equation

$$(x_1+x_2)\frac{\partial G}{\partial x_0} + (x_0+x_2)\frac{\partial G}{\partial x_1} + (x_0+x_1)\frac{\partial G}{\partial x_2} - \frac{\partial G}{\partial t} = 2NG$$

The solution of this first order equation is easily obtained by the characteristic equation method

$$
\begin{aligned}
& G(x_0, x_1, x_2\,;\,t) \\
& = (\tfrac{1}{3})^N \sum_{\substack{n_0, n_1, n_2 \\ m_0+m_1+m_2=N}} [(2\exp(-3t)+1)x_0 + (-\exp(-3t)+1)x_1 \\
& \qquad + (-\exp(-3t)+1)x_2]^{n_0} \\
& \quad \times [(-\exp(-3t)+1)x_0 + (2\exp(-3t)+1)x_1 + (-\exp(-3t)+1)x_2]^{n_1} \\
& \quad \times [(\exp(-3t)+1)x_0 + (-\exp(-3t)+1)x_1 \\
& \qquad + (2\exp(-3t)+1)x_2]^{n_2} \cdot p(n_0, n_1, n_2\,;\,t=0)
\end{aligned}
$$

This result gives the evolution to equilibrium for a system of N tree valued spins starting from an arbitrary initial state. As expected, the equilibrium state is simply the trinomial distribution.

3. The general relation between the evolution of a symmetric system of two valued elements and the birth and death equations

The discussion in § 2.2.1 makes it clear that the Liouville equation of a symmetric system of two valued elements is quite generally a birth and death equation of the form (Bailey 1964, p. 101)

$$\dot{p}_n = -(\lambda_n + \mu_n)p_n + \lambda_{n-1}p_{n-1} + \mu_{n+1}p_{n+1}$$

with

$$\left.\begin{aligned} \lambda_n &= \frac{N-n}{n+1}\sum_{i=1}^{n+1} w_i(\boldsymbol{\sigma}_i) \\ \mu_n &= \frac{n}{N-n+1}\sum_{i=n}^{N} w_i(\boldsymbol{\sigma}_i) \end{aligned}\right\} \tag{6}$$

What we are going to show now is that the birth and death rates λ_n, μ_n are : linear for independent elements (possibly with an external interaction) ; quadratic for binary interactions ; and cubic for third-order interactions.

3.1. *Preliminaries*

First of all, it is clear from the relations (6) defining λ_n and μ_n that the birth and death transition rates for the process

$$w_i(\boldsymbol{\sigma}) = w_i^{(1)}(\boldsymbol{\sigma}) + w_i^{(2)}(\boldsymbol{\sigma})$$

are

$$\lambda_n = \lambda_n^{(1)} + \lambda_n^{(2)}$$

$$\mu_n = \mu_n^{(1)} + \mu_n^{(2)}$$

where $\lambda_n^{(k)}$, $\mu_n^{(k)}$, $k=1, 2$ are the transition rates corresponding to the process $w_i^{(k)}(\boldsymbol{\sigma})$.

Let us consider an interaction term of a fairly general form

$$w_i(\boldsymbol{\sigma}) = a + u(\sigma_i)U(\boldsymbol{\sigma}'_i)$$

with

$$\boldsymbol{\sigma}'_i = (\sigma_1, \ldots, \sigma_{i-1}, \sigma_{i+1}, \ldots, \sigma_N)$$

and U being a symmetrical function of its arguments. The positive number a has been introduced merely to secure the positiveness of $w_i(\boldsymbol{\sigma})$.

First we have to check the symmetry conservation property. There are two kinds of terms

$$\sum_{i=1}^{N} w_i(\boldsymbol{\sigma}) = \sum_{i=1}^{N} u(\sigma_i)U(\boldsymbol{\sigma}'_i) \tag{7}$$

$$\sum_{i=1}^{N} w_i(\boldsymbol{\sigma}_i)p(\boldsymbol{\sigma}_i\,;\,t) = \sum_{i=1}^{N} u(-\sigma_i)U(\boldsymbol{\sigma}'_i)p(\boldsymbol{\sigma}_i\,;\,t) \tag{8}$$

Inverting σ_j and σ_k produces two kinds of modifications :

(i) The terms $i=j, k$, that is $u(\sigma_j)U(\boldsymbol{\sigma}'_j)$, $u(\sigma_k)U(\boldsymbol{\sigma}'_k)$ in eqn. (7), and $u(-\sigma_j)U(\boldsymbol{\sigma}'_j)p(\boldsymbol{\sigma}_j\,;\,t)$, $u(-\sigma_k)U(\boldsymbol{\sigma}'_k)p(\boldsymbol{\sigma}_k\,;\,t)$ in eqn. (8) are simply interchanged.

(ii) Moreover due to the symmetry of the function U, the other terms remain equal to themselves in the interchange of σ_j and σ_k.

The corresponding birth and death transition rates are obtained in the same way as in § 2.2.1. This yields

$$\left.\begin{aligned}\lambda_n &= (N-n)[a+u(-1)U(n)]\\ \mu_n &= n[a+u(1)U(n-1)]\end{aligned}\right\} \quad (9)$$

where $U(n)$ has the following meaning

$$U(n) = U(\underbrace{1, 1, \ldots, 1}_{n \text{ times}}, \underbrace{-1, \ldots, -1}_{\substack{N-n-1\\ \text{times}}})$$

3.2. *Independent particles with external interaction*

This case corresponds to the following transition rate for the ith particle

$$w_i(\boldsymbol{\sigma}) = u(\sigma_i)$$

that is the transition probabilities of a particle depends only on its own state : this is typical of a system subject only to an external interaction. It is already clear from eqn. (9) (with $U=1$) that the corresponding birth and death process will be a linear one ; actually it is the most general finite process admitting the reflecting states $n=0$ and $n=N$.

3.3. *Binary interaction*

The transition rate for the ith particle is now

$$w_i(\boldsymbol{\sigma}) = \sum_{j \neq i} v(\sigma_i, \sigma_j)$$

The function $v(\sigma, \sigma')$ may be uniquely represented as

$$v(\sigma, \sigma') = a + b\sigma + c\sigma' + d\sigma\sigma'$$

The choice of a, b, c, d must of course, preserve the positiveness of $v(\sigma, \sigma')$; in particular a must be positive. The term $b\sigma$ was considered in the previous case.

Before proceeding, let us give the interpretation of the term $a+d\sigma\sigma'$ taken alone. If $d<0$ (and $a>|d|$), we get for $v(\sigma, \sigma') = a + d\sigma\sigma'$

$$v(1, 1) < v(1, -1)$$

and

$$v(1, 1) = v(-1, -1)$$

$$v(1, -1) = v(-1, 1)$$

A particle is thus less likely having a transition if its neighbours are in the same state : this is an imitative behaviour. In physical terms, this interaction favours the parallel configurations of spins ; this results in a ferromagnetic tendency of the system. To model the spreading of an epidemic for example, we could adopt this $v(\sigma, \sigma')$, with $\sigma=-1$ meaning illness. However

$$v(-1, 1) < v(1, -1)$$

seems more reasonable. This can be realized by adding to $v(\sigma, \sigma')$ a term

$$b(\sigma - \sigma'), \quad b > 0$$

This indicates the possible role of the term $c\sigma'$ which does not have any interesting interpretation if taken alone.

If $d > 0$, the contrary happens : it is a behaviour with a tendency to contradiction. In physical terms, this interaction favours the antiparallel configurations, thus implying a tendency to antiferromagnetism.

An easy calculation shows that the corresponding birth and death process is the most general finite, quadratic process admitting $n=0$ and $n=N$ as reflecting states

$$\left.\begin{aligned} \lambda_n &= (N-n)(\alpha n+\beta) \\ \mu_n &= n(\tilde{\alpha} n+\tilde{\beta}) \end{aligned}\right\} \tag{10}$$

with

$$\left.\begin{aligned} \alpha &= 2(c-d), \quad \beta=(N-1)(a-b-c+d) \\ \tilde{\alpha} &= 2(c+d), \quad \tilde{\beta}=(N-1)(a+b)-(N+1)(c+d) \end{aligned}\right\} \tag{11}$$

The solution of this process is known (Karlin and McGregor 1962) if the λ_n and μ_n are asymptotically symmetric, that is if

$$\alpha=-\tilde{\alpha} \quad \text{and} \quad \tilde{\alpha}>0$$

which results in the condition

$$c=0$$

It will be shown in § 3.4.3 that this case corresponds precisely to a reduced distribution function hierarchy which is a closed one. The equilibrium distribution of the general process (10) has a simple analytical form which will be considered later in the paper.

The particular case studied by Kac (1977) is

$$\lambda_n=(N-n)\,\frac{N-n-1}{N}$$

$$\mu_n=n\,\frac{N-n}{N}$$

which corresponds to the following $v(\sigma, \sigma')$

$$v(\sigma, \sigma')=1-\frac{1}{2(N-1)}+\frac{\sigma}{2N(N-1)}-\frac{\sigma'}{2N}$$

Kac's topic was rather the illustration of statistical mechanic's ideas. No attempt was made to obtain the exact time dependent solution of this process.

3.4. *Triple interaction*

The transition rate for the ith particle is

$$w_i(\boldsymbol{\sigma})=\sum_{\substack{j\neq i\\ k\neq i\\ j\neq k}} v(\sigma_i, \sigma_j, \sigma_k)$$

The decomposition of the function $v(\sigma, \sigma', \sigma'')$ now reads

$$v(\sigma, \sigma', \sigma'')=a+b_1\sigma_1+b_2\sigma_2+b_3\sigma_3+c_1\sigma_2\sigma_3+c_2\sigma_1\sigma_3+c_3\sigma_1\sigma_2+d\sigma_1\sigma_2\sigma_3 \tag{12}$$

There is still a one to one relation between the values of the function v and the coefficients a, b, c, but this time the proof involves an eighth-order determinant. Let us show that it is indeed different from zero. First we notice that

$$a = \frac{1}{8} \sum_{\sigma, \sigma', \sigma''} v(\sigma, \sigma', \sigma'')$$

Thus we need only seven equations. We disregard the equation corresponding to $(\sigma, \sigma', \sigma'') = (1, 1, 1)$ and we denote by A the matrix of this system. It happens that the product of A by its transpose is a cyclical matrix. The corresponding determinant is equal, except of a sign factor, to (Kowalewski 1960, p. 105)

$$\prod_{i=1}^{7} f(x_i)$$

where $x_1, \ldots, x_7$ are the seven roots of : $x^7 = 1$, and

$$f(x) = 7 - (x + \ldots + x^6)$$

Except for $x_1 = 1$, we have

$$\sum_{k=0}^{6} x_i^k = 0 \Rightarrow f(x_i) = 8$$

for $x_1 = 1$, we have $f(x_1) = 1$. Thus

$$\det (A\tilde{A}) = \det^2 (A) \neq 0$$

For brevity, we will not consider the transition rates λ_n, μ_n in the general case. We shall be content with the examination of the term

$$v(\sigma, \sigma', \sigma'') = a + d\sigma\sigma'\sigma''$$

the other terms being very similar. This term is of the general form

$$w_i(\boldsymbol{\sigma}) = \alpha + u(\sigma_i) U(\boldsymbol{\sigma}'_i)$$

with

$$U(\boldsymbol{\sigma}'_i) = \sum_{\substack{1 \leq j < k \leq N \\ j, k \neq i}} \sigma_j \sigma_k$$

Thus the symmetry conservation is automatically satisfied. We have now to evaluate the function : $U(n)$. In the expression of $U(n)$, there are n numbers equal to one and $N-1-n$ numbers equal to minus-one. Thus

$$U(n) = \binom{n}{2} + \binom{N-1-n}{2} - n(N-1-n)$$

Applying formula (9) yields

$$\lambda_n = (N-n) \left[-2dn^2 + 2(N-1)dn + \frac{(N-1)(N-2)}{2} (a-d) \right]$$

$$\mu_n = n \left[2dn^2 - 2(N+1)dn + \frac{N-2}{2} (a(N-2) + d(N+2)) \right]$$

This is a cubic asymptotically symmetric process.

In the general case, we shall obtain the most general cubic process admitting the reflecting states : $n=0$ and $n=N$. However, as we already noticed, binary interactions seem general enough to represent the interactions occurring in nature.

4. The equations for the mean values and the reduced distribution functions

In non-equilibrium statistical mechanics, starting from the Liouville equation one derives the equations for the reduced distribution functions

$$F^{(p)}(x_1, \ldots, x_p\ ;\ t)=N(N-1)\ldots(N-p+1)\int F(x_1, \ldots, x_N\ ;\ t)\, dx_{p+1}\ldots dx_N$$

One could as well introduce the p-particles means

$$\underset{1,2,\ldots,p}{m}(t)=\langle x_1\ldots x_p\rangle=\int x_1\ldots x_p\, F(x_1, \ldots, x_N\ ;\ t)\, dx_1\ldots dx_N$$

In our case, these two sets of functions are completely equivalent, as will be shown below, for every integer p.

The equations for the one and two particle means have been given by Glauber (1963). For the sake of completeness, we shall briefly recall them here, along with the equations for the reduced distribution functions.

4.1. *Equations for the one and two particles means*

The one and two particles means are here defined as

$$m_i(t)=\sum_{\boldsymbol{\sigma}}\sigma_i p(\boldsymbol{\sigma}\ ;\ t), \quad 1\leqslant i\leqslant N$$

$$m_{ij}(t)=\sum_{\boldsymbol{\sigma}}\sigma_i\sigma_j p(\boldsymbol{\sigma}\ ;\ t), \quad 1\leqslant i<j\leqslant N$$

where the summation is over all the values of $\sigma_1, \ldots, \sigma_N$. Due to the symmetry of $p(\boldsymbol{\sigma}\ ;\ t)$ all the one (respectively two) particle means will be equal

$$m_i(t)=m_1(t)$$

$$m_{ij}(t)=m_{12}(t)$$

To obtain the equation for $m_1(t)$, we start from the Kolmogorov equation (1). We multiply both sides by σ_1 and sum over $\sigma_1, \ldots, \sigma_N$. We get thus

$$\sum_{\boldsymbol{\sigma}}\sigma_1\dot{p}(\boldsymbol{\sigma}\ ;\ t)=-\sum_{\boldsymbol{\sigma}}\sigma_1 w_1(\boldsymbol{\sigma})p(\boldsymbol{\sigma}\ ;\ t)-\sum_{\boldsymbol{\sigma}}\sigma_1\sum_{i=2}^{N}w_i(\boldsymbol{\sigma})p(\boldsymbol{\sigma}\ ;\ t)$$

$$+\sum_{\boldsymbol{\sigma}}\sigma_1 w_1(\boldsymbol{\sigma}_1)p(\boldsymbol{\sigma}_1\ ;\ t)+\sum_{\boldsymbol{\sigma}}\sigma_1\sum_{i=2}^{N}w_i(\boldsymbol{\sigma}_i)p(\boldsymbol{\sigma}_i\ ;\ t)$$

The second terms in each row cancel out and the first one doubles. Thus

$$\dot{m}_1=(-2)\langle S_1\cdot w_1(\mathbf{S})\rangle$$

where the capital letters denote the random variables and $\langle\,\cdot\,\rangle$ the expectation.

In the same way, we get the equation for the two particles means

$$\dot{m}_{12}=(-2)\langle S_1\cdot S_2(w_1(\mathbf{S})+w_2(\mathbf{S}))\rangle \tag{13}$$

4.2. *Equation for the reduced distribution function*

The reduced distribution function of order one and p are defined as

$$f_i(\sigma_i\,;\ t) = \sum_{\boldsymbol{\sigma}'_i} p(\boldsymbol{\sigma}\,;\ t)$$

$$f_{i_1,\dots,i_p}(\sigma_{i_1},\dots,\sigma_{i_p}\,;\ t) = \sum_{\boldsymbol{\sigma}'_{i_1,\dots,i_p}} p(\boldsymbol{\sigma}\,;\ t)$$

where

$$\boldsymbol{\sigma}'_{i_1,\dots,i_p} = (\sigma_1,\ \dots,\ \sigma_N) - (\sigma_{i_1},\dots,\sigma_{i_p})$$

Of course, all those functions of the same order are again equal. Hence it is natural to define the one and p order reduced distribution function as

$$f_1(\sigma_1\,;\ t) = N \sum_{\boldsymbol{\sigma}'_1} p(\boldsymbol{\sigma}\,;\ t)$$

$$f_p(\sigma^{(p)}\,;\ t) = N(N-1)\ \dots\ (N-p+1) \sum_{\boldsymbol{\sigma}'_{1,2,\dots,p}} p(\boldsymbol{\sigma}\,;\ t)$$

where

$$\sigma^{(p)} = (\sigma_1,\ \sigma_2,\ \dots,\ \sigma_p)$$

In particular the function $f_2(\sigma^{(2)}\,;\ t)$ is closely related to the pair correlation function (Reed and Gubbins 1973, p. 175) for which direct measurements are possible.

Obtaining the evolution equation for $f_p(\sigma^{(p)}\,;\ t)$ uses the same procedure as before and we get

$$\frac{\dot{f}_p(\sigma^{(p)}\,;\ t)}{N(N-1)\ \dots\ (N-p+1)} = -p(\boldsymbol{\sigma}\,;\ t) \sum_{i=1}^{p} \sum_{\boldsymbol{\sigma}'_{1,2,\dots,p}} w_i(\boldsymbol{\sigma}) + \sum_{i=1}^{p} \sum_{\boldsymbol{\sigma}'_{1,2,\dots,p}} w_i(\boldsymbol{\sigma}_i) p(\boldsymbol{\sigma}_i\,;\ t) \tag{14}$$

To give that equation a more suggestive form, we have to choose a specific form of the interaction. Let us take a binary interaction

$$w_i(\boldsymbol{\sigma}) = \sum_{j\neq i} v(\sigma_i,\ \sigma_j)$$

We obtain thus

$$\dot{f}_p(\sigma^{(p)}\,;\ t) = -f_p(\sigma^{(p)}\,;\ t) \sum_{i=1}^{p} h_i(\sigma^{(p)}) + \sum_{i=1}^{p} f_p(\sigma_i{}^{(p)}\,;\ t) h_i(\sigma_i{}^{(p)}) + \sum_{\sigma} \left[-f_{p+1}(\sigma^{(p)},\ \sigma\,;\ t) \sum_{i=1}^{p} v(\sigma_i,\ \sigma) + \sum_{i=1}^{p} f_{p+1}(\sigma_i{}^{(p)},\ \sigma\,;\ t) v(-\sigma_i,\ \sigma) \right] \tag{15}$$

The derivation is given in Appendix 2 along with the definition of the functions h_i. Thus we get an equations hierarchy for the functions f_i with a structure very similar to the BBGKY hierarchy of statistical mechanics (Balescu 1975, p. 80), the usual integral over the phase space variable in the right hand side being replaced by a summation over σ which is the phase variable of our particles.

Before closing these general considerations let us recall that the p-particles means and the reduced distribution function are tightly related in our case.

Since the function $\frac{1}{2}(1+\sigma\sigma')$ equals unity for $\sigma=\sigma'$ and zero for $\sigma=-\sigma'$, $p(\boldsymbol{\sigma}\,;\,t)$ may be written

$$p(\boldsymbol{\sigma}\,;\,t)=\frac{1}{2^N}\sum_{\boldsymbol{\sigma}'}(1+\sigma_1\sigma'_1)\,\dots\,(1+\sigma_N\sigma'_N)p(\boldsymbol{\sigma}'\,;\,t)$$

$$=\frac{1}{2^N}\left[1+\sum_k\sigma_k m_k(t)+\sum_{\substack{k,l\\k\neq l}}\sigma_k\sigma_l m_{kl}(t)+\dots\right]$$

Using this expression of $p(\boldsymbol{\sigma}\,;\,t)$ in the definitions of the reduced distribution function, we get

$$f_i(\sigma_i\,;\,t)=\frac{1}{2^N}\left[\sum_{\boldsymbol{\sigma}'_i}1+\sigma_i m_i(t)\sum_{\boldsymbol{\sigma}'_i}1+\sum_{k\neq i}m_k(t)\sum_{\sigma_k}\sigma_k\sum_{\boldsymbol{\sigma}'_{i,k}}1\right]$$

that is

$$f_i(\sigma_i\,;\,t)=\tfrac{1}{2}(1+\sigma_i m_i(t))$$

and in the same way

$$f_{ij}(\sigma_i,\sigma_j\,;\,t)=\tfrac{1}{4}(1+\sigma_i m_i(t)+\sigma_j m_j(t)+\sigma_i\sigma_j m_{ij}(t))$$

4.3. *Particular cases*

It is interesting to find out for which binary interactions the hierarchy for the reduced distribution function closes. Let us first examine eqn. (15) for $p=1$. Replacing $v(\sigma,\sigma')$ by

$$v(\sigma,\sigma')=a+b\sigma+c\sigma'+d\sigma\sigma'$$

the right hand side becomes

$$a\sum_{\sigma}(-f_2(\sigma_1,\sigma)+f_2(-\sigma_1,\sigma))-b\sigma_1\sum_{\sigma}(f_2(\sigma_1,\sigma)+f_2(-\sigma_1,\sigma))$$
$$+c\sum_{\sigma}\sigma(-f_2(\sigma_1,\sigma)+f_2(-\sigma_1,\sigma))-d\sigma_1\sum_{\sigma}\sigma(f_2(\sigma_1,\sigma)+f_2(-\sigma_1,\sigma))$$

All terms reduced to f_1, except the third. For the equation to be free of f_2 requires $c=0$. We now have the general result expressed in the following proposition.

Proposition 1

The necessary and sufficient condition for the reduced distributions hierarchy to close at all orders in the case of a binary interaction is that

$$v(\sigma,\sigma')=a+b\sigma+d\sigma\sigma'$$

The proof is given in Appendix 2.

We observed in § 2.3.3 that the only solvable (up to now!) binary interaction model is precisely that for which $c=0$. This is not surprising in the light of the present property. There is a related property for the corresponding birth and death process which is a quadratic asymptotically symmetric one. Namely, it is the only birth and death process for which the set of evolution equations for the moments

$$m_k(t) = \sum_n n^k p_n(t)$$

closes (Roehner and Valent 1982).

In his paper, Glauber (1963) studied the interaction between nearest neighbours. It is straightforward to observe that the symmetry conservation property does not hold in that case. This means that starting with a symmetric function for $p(\boldsymbol{\sigma}\,;\ t=0)$, $p(\boldsymbol{\sigma}\,;\ t)$ does not remain symmetric. This looks at first sight surprising since all elements play individually the same role (there is no boundary effect for the cyclical condition is assumed at the end of the chain). However, the actual reason becomes clear if we look at the equations for the p-particles means. For nearest neighbours interaction (Glauber 1963) eqn. (13) for the two particles means becomes

$$\left.\begin{aligned} \dot{m}_{ij} &= -2m_{ij} + \tfrac{1}{2}\gamma(m_{i,j-1} + m_{i,j+1} + m_{i-1,j} + m_{i+1,j}) \\ m_{ii} &\equiv 1, \quad i \neq j \end{aligned}\right\} \tag{16}$$

Equation (16) obviously takes two different forms depending on the value of $|i-j|$. Indeed if $|i-j|=1$, there are m_{ii} terms (reducing to unity) which is not the case if $|i-j| \geqslant 2$.

Thus, even if all $m_{ij}(t)$ are equal initially (as a consequence of $p(\boldsymbol{\sigma}\,;\ t=0)$ being symmetric) they will not remain equal. This is of course, directly related to the kind of interaction considered. In § 5.3.2 the consequence of this fact in relation to the postulate of equal *a priori* probabilities will be discussed.

5. Qualitative features of the system's evolution

Our purpose so far has mainly been to emphasize that this system of interacting elements bears a close relation to the classical birth and death equations. However, since it is mathematically almost the simplest system involving strongly interacting elements, it is tempting to take advantage of that simplicity to study the general features of the system's evolution in time. Although we shall mainly use the language of statistical mechanics, it should be remembered that such considerations could apply to other fields too. The identification of possible constants of the motion is the first important question.

5.1. *Existence of constants of the motion*

Since our system is a stochastic one, a possible constant of the motion $h(\boldsymbol{\sigma})$ can only be conserved in the mean, that is

$$\sum_{\boldsymbol{\sigma}} h(\boldsymbol{\sigma})p(\boldsymbol{\sigma}\,;\ t) = \text{constant}$$

Due to the symmetry of $p(\boldsymbol{\sigma}\,;\ t)$ with respect to the σ_i, $h(\boldsymbol{\sigma})$ should be a symmetric function too. The conservation condition reads thus

$$\sum_{\boldsymbol{\sigma}} h(n)\left[p_n(t)\Big/\binom{N}{n}\right] = \sum_{n=0}^{N} h(n)p_n(t) = \text{constant}$$

The equation of motion of the system is the general birth and death equation

$$\dot{p}_n = -(\lambda_n + \mu_n)p_n + \lambda_{n-1}p_{n-1} + \mu_{n+1}p_{n+1} \tag{17}$$

with

$$\lambda_N = \mu_0 = 0$$

going to the derivative in the conservation condition gives, after a few arrangements

$$\sum_{n=0}^{N} (\lambda_n \phi_n - \mu_n \phi_{n-1})p_n = 0$$

where

$$\phi_n = h(n+1) - h(n)$$

Equating all the coefficients of the p_n to zero will at least give a sufficient conservation condition. However, without additional requirements it yields but a trivial conservation property. Indeed, $\mu_0 = 0$ leads to $\phi_n = 0$ for all n; the corresponding conservation relation is thus simply the normalization of the distribution p_n.

There are (at least) two ways to avoid this difficulty: (i) by admitting $n=0$ as an absorbing state; and (ii) by a mirror process. Let us consider the first way.

We assume that $\lambda_0 = \mu_0 = 0$ and $N = \infty$. The conservation condition becomes

$$\sum_{n=1}^{\infty} (\lambda_n \phi_n - \mu_n \phi_{n-1})p_n = 0$$

It is fulfilled if

$$\phi_n = \frac{\mu_1 \dots \mu_n}{\lambda_1 \dots \lambda_n} \phi_1, \quad n \geqslant 1$$

Solving the recurrence relation for $h(n)$ gives

$$\left.\begin{aligned} h(n) &= a + b + b \sum_{k=1}^{n-1} \frac{\mu_1 \dots \mu_k}{\lambda_1 \dots \lambda_k}, \quad n \geqslant 2 \\ h(1) &= a + b \\ h(0) &= a \end{aligned}\right\} \tag{18}$$

where a and b are arbitrary constants.

Needless to say, our derivation was rather formal and for the result to be meaningful, the series $\sum_{n=0}^{\infty} h(n)p_n(t)$ must converge. This is the case in the following example.

Example

For the infinite process studied in § 2.2.2, we have

$$\lambda_n = \lambda \, . \, n$$

$$\mu_n = \mu \, . \, n$$

which give

$$h(n) = a + b + b\,\frac{\rho^n - \rho}{\rho - 1}, \quad n \geqslant 0, \quad \rho = \frac{\mu}{\lambda}$$

For the sake of simplicity, take $a = 1$ and $b = \rho - 1$, then

$$h(n) = \rho^n, \quad n \geqslant 0$$

Let us check directly that $h(n)$ is conserved using the known expressions for $p_n(t)$ (Bailey 1964, p. 94) ; for an initial state $n_0=1$

$$p_n(t)=(1-\rho\beta)(1-\beta)\beta^{n-1}, \quad n\geqslant 1$$

$$p_0(t)=\rho\beta$$

β is a given function of time, but we do not need its detailed expression. Now the constant of the motion becomes

$$\sum_{n=0}^{\infty} h(n)p_n(t)=\rho\beta+(1-\rho\beta)(1-\beta)\rho \sum_{n=0}^{\infty} (\rho\beta)^n$$
$$=\rho$$

provided $\rho\beta<1$, which is true for $t<\infty$.

Now turning to mirror processes. We assume the process to be a finite one, with the usual condition

$$\mu_0=\lambda_N=0 .$$

We regroup now the terms $n=0$ and $n=N$

$$\lambda_0\phi_0 p_0-\mu_N\phi_{N-1}p_N \tag{19}$$

and we want both terms to cancel out. This can be done easily if $p_0(t)=p_N(t)$ for all t, which is so if

$$\lambda_n=\mu_{N-n}, \quad n\geqslant 0 \tag{20}$$

Actually, eqn. (20) implies more generally that (Appendix 3) : $p_n(t)=p_{N-n}(t)$, $n\geqslant 0$, $t>0$. A finite process such that : $\lambda_n=\mu_{N-n}$, $n\geqslant 0$ will be referred to as a *mirror process*.

Now the terms in eqn. (19) vanish provided $\phi_{N-1}=\phi_0$. The other terms may be annulled as before. This gives

$$\phi_n=\frac{\mu_1 \ldots \mu_n}{\lambda_1 \ldots \lambda_n}\phi_0, \quad n\geqslant 1$$

and we notice that the condition $\phi_{N-1}=\phi_0$ is automatically satisfied.

The constant of the motion (in the mean) again has the form of eqn. (18) but this time with the constraint $\lambda_n=\mu_{N-n}$. An example is provided in Appendix 3.

Let us now identify the quadratic and cubic mirror processes. The condition (20) leads immediately to the following families of processes

$$\lambda_n=(N-n)(\alpha n+\beta), \quad \lambda_n=(N-n)(\alpha n^2+\beta n+\gamma)$$

$$\mu_n=n[\alpha(N-n)+\beta], \quad \mu_n=n[\alpha(N-n)^2+\beta(N-n)+\gamma]$$

Notice that a mirror process of even degree will be an asymptotically symmetric process (Roehner and Valent 1982) while the contrary is true for an odd degree process.

To conclude this section, let us state the corresponding result for a system with binary interactions. Using eqns. (11) we derived the following proposition.

Proposition 2

For a binary interaction defined by

$$v(\sigma, \sigma') = \alpha - \beta\sigma + (N-1)\beta\sigma\sigma'$$

(where α and β should be chosen so as to secure positivity of $v(\sigma, \sigma')$). The functions $h(\sigma_1 + \ldots + \sigma_N)$ are conserved in the mean, where

$$h(s) = a + b + b \sum_{k=1}^{(s+N)/2-1} \frac{\mu_1 \ldots \mu_k}{\lambda_1 \ldots \lambda_k}, \quad s = -N+4, \; -N+6, \; \ldots, \; N$$

$$h(-N+2) = a + b$$

$$h(-N) = a$$

5.2. *Condition for irreversibility of the system's motion*

The question of the existence of a preferred direction of motion is another important qualitative property of the system's evolution in time. In physics, irreversibility is defined through the property of the entropy function to be steadily increasing. We shall follow this way of thought too. However, it should be kept in mind that this might not be the only possible approach.

The entropy of the system under consideration is

$$S(t) = -\sum_{\boldsymbol{\sigma}} p(\boldsymbol{\sigma};\; t) \ln p(\boldsymbol{\sigma};\; t)$$

With the symmetry assumption for $p(\boldsymbol{\sigma};\; t)$ we get

$$S(t) = -\sum_{n=0}^{N} p_n(t) \ln \left[p_n(t) \Big/ \binom{N}{n} \right] \tag{21}$$

To avoid confusion, when dealing with the distribution $p_n(t)$ we shall refer to eqn. (21) as the associate entropy, which is the entropy of $p(\boldsymbol{\sigma};\; t)$ but not of $p_n(t)$.

We are now in a position to study the time evolution of the entropy, not from an approximate equation such as the Boltzmann equation but from the exact (stochastic) equation of motion (17). We shall state and prove a proposition.

Proposition 3

The necessary and sufficient condition for the entropy of the system to be steadily increasing is that

$$\lambda_n = (N-n)\rho(n+1)$$

$$\mu_n = n\rho(n)$$

where $\rho(n)$ is a positive function. In the case of a binary interaction $v(\sigma, \sigma')$, this condition corresponds to the following interaction

$$v(\sigma, \sigma') = a + b\sigma'$$

Proof

We first prove that the given form of λ_n, μ_n is necessary. Assume an increasing entropy. Since any initial condition may be chosen, the stationary distribution should be the distribution for which the associate entropy is maximum. To localize this maximum, we write eqn. (21) as

$$S = -\sum_{n=1}^{N} p_n \ln p_n - p_0 \ln p_0 + \sum_{n=1}^{N} p_n \ln \binom{N}{n}$$

with

$$\frac{\partial p_0}{\partial p_n} = -1, \quad n = 1, \ldots, N$$

Hence

$$\frac{\partial S}{\partial p_n} = -\ln p_n + \ln p_0 + \ln \binom{N}{n} = 0 \Rightarrow p_n = \frac{1}{2^N}\binom{N}{n}$$

This extremum is indeed a maximum if the second derivatives satisfy the following inequalities

$$\frac{\partial^2 S}{\partial {p_1}^2} < 0, \quad \begin{vmatrix} \dfrac{\partial^2 S}{\partial {p_1}^2} & \dfrac{\partial^2 S}{\partial p_1 \, \partial p_2} \\ \dfrac{\partial^2 S}{\partial p_2 \, \partial p_1} & \dfrac{\partial^2 S}{\partial {p_2}^2} \end{vmatrix} > 0, \ldots$$

This can be easily checked using the following formula (Muir 1960, p. 142)

$$\begin{vmatrix} 1+a_1 & 1 & \ldots & 1 \\ 1 & 1+a_2 & \ldots & 1 \\ \vdots & & & \\ 1 & & \ldots & 1+a_n \end{vmatrix} = a_1 \ldots a_n \left(1 + \frac{1}{a_1} + \ldots + \frac{1}{a_n}\right)$$

We get thus

$$\begin{vmatrix} \dfrac{\partial^2 S}{\partial {p_1}^2} & \dfrac{\partial^2 S}{\partial p_1 \, \partial p_2} & \ldots & \dfrac{\partial^2 S}{\partial p_1 \, \partial p_n} \\ \vdots & & & \\ \dfrac{\partial^2 S}{\partial p_n \, \partial p_1} & & \ldots & \dfrac{\partial^2 S}{\partial {p_n}^2} \end{vmatrix} = \left(-\frac{1}{p_0}\right)^n \frac{1}{\prod_{i=1}^{n} \binom{N}{i}} \sum_{i=0}^{n} \binom{N}{i}$$

Let us now identify the stationary distribution (Parzen 1962, p. 280)

$$p_n(\infty) = \frac{\lambda_0, \ldots, \lambda_{n-1}}{\mu_1, \ldots, \mu_n} p_0(\infty)$$

with the binomial distribution that maximizes S. This leads immediately to

$$\left.\begin{aligned} \lambda_n &= (N-n)\rho(n+1) \\ \mu_n &= n\rho(n) \end{aligned}\right\} \qquad (22)$$

where $\rho(n)$ is any positive function.

To prove that this condition is also sufficient, we shall show that it implies

$$\dot{S}(t) \geqslant 0$$

Due to the normalization condition for the p_n

$$\dot{S}(t) = -\sum_{n=0}^{N} \dot{p}_n \ln \left[p_n \Big/ \binom{N}{n} \right]$$

Taking into account the equations of motion (17)

$$\dot{S}(t)=\sum_{n=0}^{N}\left[(\lambda_n+\mu_n)p_n\ln\left[p_n\Big/\binom{N}{n}\right]\right.$$
$$\left.-\lambda_n p_n\ln\left[p_{n+1}\Big/\binom{N}{n+1}\right]-\mu_n p_n\ln\left[p_{n-1}\Big/\binom{N}{n-1}\right]\right]$$

where two obvious index changes have been carried out. We regroup the λ_n and the μ_n terms

$$\dot{S}(t)=\sum_{n=0}^{N}\left[\lambda_n p_n\ln\frac{p_n}{p_{n+1}}\frac{N-n}{n+1}+\mu_n p_n\ln\frac{p_n}{p_{n-1}}\frac{n}{N-(n-1)}\right]$$

We use now the inequality

$$\ln x\geqslant 1-\frac{1}{x},\quad x>0$$

along with the specific form of λ_n and μ_n given by eqn. (22)

$$\dot{S}(t)\geqslant\sum_{n=0}^{N}[(N-n)\rho(n+1)p_n-\rho(n+1)(n+1)p_{n+1}+n\rho(n)p_n$$
$$-\rho(n)(N-(n-1))p_{n-1}]$$

It is clear that the right hand side vanishes and hence $\dot{S}(t)\geqslant 0$.

To get the last part of Proposition 3, we have to use eqns. (11). □

5.3. *Discussion of the postulate of equal* a priori *probabilities*

The postulate of equal *a priori* probabilities for all states of an isolated system is the basis of equilibrium statistical mechanics. It is also invoked in fields other than physics (Bussières 1982, p. 47).

Our N particles system is simple enough for the exact equilibrium distribution to be written out explicitly, at least for a binary interaction. Hence, it is interesting to see which interactions lead to an equilibrium distribution compatible with that postulate.

5.3.1. *The stationary distribution*

We restrict ourselves to the binary interaction case which is most interesting and for which the λ_n, μ_n are given by eqns. (10). The stationary condition results for the generating function

$$\phi(x)=\sum_{n=0}^{N}x^n p_n(\infty)$$

in the differential equation

$$\tilde{\alpha}x\left(1+\frac{\alpha}{\tilde{\alpha}}x\right)\phi''+[\tilde{\alpha}+\tilde{\beta}-x((N-1)\alpha+\beta)]\phi'-N\beta\phi=0$$

which is, apart from a variable change, an hypergeometric equation, the solution to which is

$$\phi(x)=\frac{{}_2F_1\left(-N,\frac{\beta}{\alpha};\ 1+\frac{\tilde{\beta}}{\tilde{\alpha}};\ -\frac{\alpha}{\tilde{\alpha}}x\right)}{{}_2F_1\left(-N,\frac{\beta}{\alpha};\ 1+\frac{\tilde{\beta}}{\tilde{\alpha}};\ -\frac{\alpha}{\tilde{\alpha}}\right)}$$

It is a polynomial of degree N.

5.3.2. *Discussion*

Proposition 4

The postulate of equal *a priori* equilibrium probabilities is satisfied for the binary interaction

$$v(\sigma,\sigma')=a+c\sigma'$$

The postulate of equal *a priori* probabilities implies that

$$p(\boldsymbol{\sigma};\ \infty)=\frac{1}{2^N}$$

that is

$$p_n(\infty)=\frac{1}{2^N}\binom{N}{n}$$

Sufficient conditions for this are (Gradshteyn 1965, p. 1040)

$$\tilde{\alpha}=\alpha,\quad \tilde{\beta}=\beta-\alpha$$

Replacing α, β, $\tilde{\alpha}$, $\tilde{\beta}$ by their expressions (11) proves the result.

Comments

Again, we have a very limited condition, the same as in Proposition 3. Actually, it seems obvious to us that the postulate of equal *a priori* probabilities cannot be true in the very strong form in which it is usually stated.

Consider for example a spin lattice (in any dimension) with nearest neighbours interaction. We observed already (§ 4.3) that, due to the form of the interaction, the two particle means $m_{ij}(t)$ behave differently whether i, j refers to pairs of nearest neighbours or not. Thus it is very unlikely for the equilibrium state to have the kind of symmetry required by the postulate of equal probabilities. This intuitive reasoning is substantiated (in one dimension) by Glauber's calculation (1963, p. 301)†, where it is found that

$$m_{ij}(t)\underset{t\to\infty}{\longrightarrow}\eta^{i-j}$$

where η is a positive number.

The kind of interaction considered in this article being more symmetric, should be favourable to the equal probabilities postulate. Nevertheless it holds in very special cases only.

---

† The example considers a system interacting with a heat reservoir, but this assumption does not affect the calculations.

Actually this is not at all a problem for statistical mechanics. The postulate which can truly be confronted with experimental facts is much less demanding, since any observation requires many states and particles to be lumped together. The problem still remains of a redefinition of this postulate.

Appendix 1

We prove that a function $F(\sigma_1, \ldots, \sigma_n)$ depending symmetrically on two valued arguments $\sigma_1, \ldots, \sigma_n$ is actually a function of the sum $s=\sigma_1+\ldots+\sigma_n$, i.e.

$$F(\sigma_1, \ldots, \sigma_n)=f(s)$$

This can be seen in two different ways.

The first way is to say the function F can assume $n+1$ values (not necessarily all distinct) denoted as $F_1, \ldots, F_{n+1}$. Now the sum s itself takes on $n+1$ different values: $s_1, \ldots, s_{n+1}$. Hence it is easy to define a function $f(s)$ such that

$$f(s_i)=F_i$$

The second reasoning is less general since we assume x_i to take the values ∓ 1, but it shows how to construct the function $f(s)$ effectively.

We assume that $F(\sigma_1, \ldots, \sigma_n)$ is a symmetric polynomial of the σ_i. It is known that such a polynomial can be expressed in terms of the fundamental functions (Fricke 1924)

$$\phi_k=\sigma_1{}^k+\ldots+\sigma_n{}^k$$

where k is a positive integer. Then, for any k

$$\phi_{2k}=n \quad \text{and} \quad \phi_{2k+1}=\sigma_1+\ldots+\sigma_n$$

The second reasoning need only be modified slightly if the σ_i take three values

$$\sigma_i=-1, 0, 1$$

Now the fundamental functions ϕ_k have two different forms. Since for $\sigma_i=-1, 0, 1$

$$\sigma_i{}^{2k+1}=\sigma_i$$

$$\sigma_i{}^{2k}=\sigma_i{}^2$$

the symmetric polynomial $F(\sigma_1, \ldots, \sigma_n)$ is actually a function of

$$s=\sigma_1+\ldots+\sigma_n$$

and

$$r=\sigma_1{}^2+\ldots+\sigma_n{}^2$$

Appendix 2

Replace

$$w_i(\boldsymbol{\sigma})=\sum_{j\neq i} v(\sigma_i, \sigma_j)$$

in eqn. (14) and multiply both sides by $N(N-1)\ldots(N-p+1)$. Two different terms then occur depending upon whether j is larger than p or not. For $j \geqslant p+1$, we get in the right hand side

$$N(N-1)\ldots(N-p+1)\sum_{i=1}^{p}\sum_{j\geqslant p+1}\sum_{\sigma_j}\left[(-v(\sigma_i,\sigma_j))\sum_{\sigma'_{1,2,\ldots,p_j}} p(\boldsymbol{\sigma};t) + v(-\sigma_i,\sigma_j)\sum_{\sigma'_{1,2,\ldots,p_j}} p(\boldsymbol{\sigma}_i;t)\right]$$

Due to the summation over σ_j, all those $N-p$ terms are equal and this gives

$$\sum_{i=1}^{p}\sum_{\sigma}[(-v(\sigma_i,\sigma_j))f_{p+1}(\sigma^{(p)},\sigma;t)+v(-\sigma_i,\sigma)f_{p+1}(\sigma_i{}^{(p)},\sigma;t)$$

with

$$\sigma_i{}^{(p)}=(\sigma_1,\ldots,-\sigma_i,\ldots,\sigma_p)$$

For $1\leqslant j\leqslant p$, $j\neq i$, we obtain in a similar way

$$-f_p(\sigma^{(p)};t)\sum_{i=1}^{p}h_i(\sigma^{(p)})+f_p(\sigma_i{}^{(p)};t)\sum_{i=1}^{p}h_i(\sigma_i{}^{(p)})$$

with

$$h_i(\sigma_1,\ldots,\sigma_p)=\sum_{\substack{1\leqslant j\leqslant p\\ j\neq i}} v(\sigma_i,\sigma_j)$$

Collecting all terms gives eqn. (15).

Assuming now that

$$v(\sigma,\sigma')=a+b\sigma+d\sigma\sigma'$$

we shall show that the right hand side involves only f_p. For the term $a+b\sigma$, f_{p+1} reduces to f_p because of the summation over σ.

For $d\sigma\sigma'$, the f_{p+1} terms have the following expression

$$(-d)\sum_{i=1}^{p}\sigma_i\sum_{\sigma}\sigma[f_{p+1}(\sigma^{(p)},\sigma;t)+f_{p+1}(\sigma_i{}^{(p)},\sigma;t)]$$

which clearly reduces to $f_p(\sigma_i{}^{(p)\prime},\sigma;t)$ where

$$\sigma_i{}^{(p)\prime}=(\sigma_1,\ldots,\sigma_{i-1},\sigma_{i+1},\ldots,\sigma_p)$$

Appendix 3

We first prove that

$$\lambda_n=\mu_{N-n},\quad n=0,1,\ldots,N \tag{A 1}$$

implies

$$p_n(t)=p_{N-n}(t),\quad n=0,1,\ldots,N,\quad t>0 \tag{A 2}$$

Assume that (A 2) is satisfied at time t, then (A 1) implies that it will hold also at time $t+\Delta t$, indeed

$$\begin{aligned}p_{N-n}(t+\Delta t)&=p_{N-n}(t)(1-\lambda_{N-n}\Delta t-\mu_{N-n}\Delta t)+\lambda_{N-(n+1)}\Delta t p_{N-(n+1)}(t)\\&\quad+\mu_{N-(n-1)}\Delta t p_{N-(n-1)}(t)\\&=p_n(t)(1-\mu_n\Delta t-\lambda_n\Delta t)+\mu_{n+1}\Delta t p_{n+1}(t)+\lambda_{n-1}\Delta t p_{n-1}(t)\end{aligned}$$

Hence, assuming that (A 2) holds at time $t=0$, it will be true for all times.

Now consider the linear mirror process

$$\left.\begin{aligned}\lambda_n &= N-n\\ \mu_n &= n\end{aligned}\right\} \qquad \text{(A 3)}$$

The function $h(n)$ becomes

$$\left.\begin{aligned}h(n) &= a+b+b\sum_{k=1}^{n-1}\left[1\Big/\binom{N-1}{k}\right]\\ h(1) &= a+b\\ h(0) &= a\end{aligned}\right\} \qquad \text{(A 4)}$$

The solution of the process (A 3) may be obtained readily by the generating function method (Bailey 1964, p. 91). This gives

$$G(x,t) \equiv \sum_{n=0}^{N} p_n(t)x^n = \left(\frac{1+x+(1-x)\exp(-2t)}{2}\right)^N \times G_0\left(\frac{1+x+(1-x)\exp(-2t)}{1+x+(1+x)\exp(-2t)}\right)$$

$$G_0(x) = G(x, t=0)$$

Assuming, for instance the mirror initial condition

$$G_0(x) = \tfrac{1}{2}(1+x^N)$$

leads to the following probabilities exhibiting the mirror property

$$p_n(t) = \left[\binom{N}{n}\Big/2^{N+1}\right](A^{N-n}B^n + A^nB^{N-n})$$

where

$$A = 1+\exp(-2t), \quad B = 1-\exp(-2t)$$

The conservation of $h(n)$ given by (A 4) may be checked directly if we notice that $h(n)+h(N-n)$ is independent of n.

References

Bailey, N. T. J., 1964, *The Elements of Stochastic Processes* (Wiley).

Balescu, R., 1975, *Equilibrium and Non-equilibrium Statistical Mechanics* (Wiley).

Bussière, R., 1982, *Systèmes Evolutifs Urbains et Regionaux à l'Etat d'Equilibre* (C.R.U.).

Fricke, R., 1924, *Lehrbuch der Algebra* (Druck und Verlag).

Glauber, R. J., 1963, *J. Math. Phys.*, **4**, 294.

Gradshteyn, I. S., and Ryzhik, I. M., 1965, *Tables of Integrals, Series and Products* (Academic Press).

Kac, M., 1977, *Statistical Mechanics and Statistical Methods in Theory and Applications*, edited by P. T. Landman (Plenum Press).

Karlin, S., and MacGregor, J., 1962, *Proc. Cambridge Phil. Soc.*, **58**, 299.

Kowalewski, G., 1960, *Einführung in die Determinantentheorie* (Chelsea).

Muir, T., 1960, *The Theory of Determinants in the Historical Order of Development* (Dover).

PARZEN, E., 1962, *Stochastic Processes* (Holden Day).
REED, T. M., and GUBBINS, K. E., 1973, *Applied Statistical Mechanics* (McGraw-Hill).
REIF, F., 1965, *Fundamentals of Statistical and Thermal Physics* (McGraw-Hill).
RICE, S. A., and GRAY, P., 1965, *The Statistical Mechanics of Simple Liquids* (Interscience).
ROEHNER, B., and VALENT, G., 1982, *SIAM Jl Appl. Math.*, **42,** 1020.
VAN KAMPEN, N. G., 1962, *Fundamental Problems in Statistical Mechanics*, edited by E. G. D. Cohen (North Holland).