Non-Markovian Model for the Approach to Equilibrium*

GÉRARD G. EMCH
University of Maryland, College Park, Maryland
(Received 1 November 1965)

This paper is intended to provide the axiomatic study of nonequilibrium quantum statistical mechanics with some simple and rigorously solvable models. The models considered here are obtained as generalizations of the Ising model. They illustrate and allow a rational discussion of the following concepts relevant to the theory of irreversible phenomena: coarse-graining and time-smoothing, ergodicity, recurrences, impossibility of a Markovian description of the approach to equilibrium for some physical systems, justification of the various random phase assumptions, properties of the interaction responsible for the approach to equilibrium, master equations, etc.

I EXPERIMENTAL BACKGROUND

THE phenomena called free-induction relaxation exhibits an oscillatory approach to equilibrium. From the first time it was observed, it was interpreted[1] as the result of the dipolar interaction between nuclear spins arranged in a rigid lattice.

Just to show that what follows is not only a mathematical game, but also a topic of some definite physical relevance, we shall recall briefly the experimental situation. This is also intended to provide a safer basis for the present theoretical considerations.

A CaF_2 crystal is placed in a magnetic field **B**, the direction of which we shall call z. The system is allowed to reach thermal equilibrium. When this is achieved, a rf pulse is applied to turn the net magnetic moment $\boldsymbol{\mu}$ in the x direction, orthogonal to the z direction of **B**. The time-evolution of the x component μ_x of $\boldsymbol{\mu}$ is then observed. It exhibits an oscillatory decay to zero, starting from a nonzero initial value. This is interpreted as the result of the dipolar interaction between the $(I = \frac{1}{2})$-spin of the fluorine nuclei. No relaxation via lattice vibration is needed to account for this phenomena, so that the spin system may be considered as isolated. It has been shown that the interaction responsible for this approach to equilibrium can be reduced to the following form (we do not want to worry here about units):

$$V = \sum_{i,j} (a_{ij}\sigma_i\sigma_j + b_{ij}\sigma_i^x\sigma_j^x) - B\sum_i \sigma_i^z, \qquad (1)$$

where $i, j \cdots$ denote the position of the fluorine nuclei, and σ^x, σ^y, σ^z are the Pauli matrices. In this form the model has not been solved exactly. Even if the following simplifying assumption does not lead to a quantitative agreement with experiment, it provides a qualitative description of the observed oscillatory approach to equilibrium, and makes the model exactly solvable:

$$a_{ij} = 0 \quad \text{for all} \quad (i, j). \qquad (2)$$

This form of the model is the basis of our considerations. The aim of these is indeed to discuss the *consequences* of the definite approach to equilibrium encountered in this problem. The remarkable fact about this model is that it does not involve any kind of repeated random phase assumption (the quantum analog of Boltzmann's *Stosszahlansatz*), and that no approximations at all are needed. The approach to equilibrium is shown to be a consequence of only the particular choice of a wide class of initial conditions (connected with the partial information obtained from macroscopic measurements) and of the peculiar form of the Hamiltonian. Because of its great simplicity, this model also allows some enlightening of many aspects of the theory of nonequilibrium processes. This is the principal motive for the present investigation.

II THEORETICAL ANALYSIS OF THE MODEL

We express the main features of our model in a quite naive mathematical form. We try to proceed in such a way that:

(i) the simplicity of the model is exhibited;
(ii) the road to the slightly more sophisticated considerations of the next section is prepared.

Consider an infinite linear chain (in several of the considerations to come it is convenient to consider the infinite linear chain as the limit of a finite ring) of fixed spin-$\frac{1}{2}$ identical particles. The quantum mechanical evolution of the system will be described by a continuous one-parameter group of unitary operators $\{U^t\}$ acting on the Hilbert space

* This research was carried out at the Institute for Fluid Dynamics and Applied Mathematics of the University of Maryland under the support of the Office of Naval Research—Contract NONR 595 (22).
[1] I. J. Lowe and R. E. Norberg, Phys. Rev. 107, 46 (1957).

$$\mathfrak{H} = \prod_i \otimes \mathfrak{H}^i, \tag{3}$$

where the index i runs over all the particles and $\mathfrak{H}^i$ are the identical replicas of the two-dimensional Hilbert space used to describe each individual particle.

We furthermore suppose that U^t can be written as

$$U^t = \prod_{n\geq 0} U_n^t, \tag{4}$$

where all the unitary groups $\{U_n^t\}$ are defined on $\mathfrak{H}$, commute among themselves, and are respectively generated by

$$\begin{aligned} H_0 &= -B \sum_i \sigma_i^z, \\ H_n &= \epsilon(n) \sum_i \sigma_i^z \sigma_{i+n}^z, \quad \text{for} \quad n > 0. \end{aligned} \tag{5}$$

We finally suppose that the real-valued function $\epsilon(n)$ satisfies the following conditions:

$$\begin{aligned} &\text{(i)} \quad \epsilon(n) \text{ is monotonically nonincreasing in } n, \\ &\text{(ii)} \quad \lim_{n\to\infty} \epsilon(n) = 0. \end{aligned} \tag{6}$$

These assumptions correspond to physical situations. For the moment we do not specify the exact form of $\epsilon(n)$, nor do we indicate whether $\epsilon(n)$ reaches its asymptotic value for finite n. Let us write for the generator of $\{U^t\}$:

$$H = \sum_{n=0}^{\infty} H_n. \tag{7}$$

Incidently, we note that

$$H^I \equiv H_0 + H_1 \tag{8}$$

is the ordinary one-dimensional Ising model with nearest-neighbor interaction only. We see that it is indeed quite essential for the approach to equilibrium to consider the generalized form of the Ising model, where the interaction extends over all pairs of neighbors.

Usually the observable of interest in the Ising model is

$$S^z = \sum_i \sigma_i^z, \tag{9}$$

which commutes with H. In this sense the ordinary Ising model, even generalized in the above way, is a classical system and is moreover only suitable for describing equilibrium situations.

Here, on the contrary, we want to consider the time-dependence of the expectation value of the observable:

$$S^z = \sum_i \sigma_i^z. \tag{10}$$

Let us now denote by $\langle S^z\rangle(0)$ the initially observed expectation value of S^z. We now have to specify the initial state of the system. As usual, many different density matrices $\rho(0)$ lead to the prescribed expectation value. If we now speak the language used, for instance, in Ref. 2, we say that these various $\rho(0)$ correspond to different microscopic states but are macroscopically equivalent. Each of them would, in principle, lead to a different time-dependence of $\langle S^z\rangle(t)$. We then have to make an assumption on the initial state we want to consider. The most reasonable choice for this is the state which maximizes the microsopic entropy, and is subject to the constraint:

$$\mathrm{Tr}\, S^z \rho(0) = \langle S^z\rangle(0). \tag{11}$$

If one takes the usual expression for the entropy, the solution of this problem is well known (it is just a transposition of one of the most satisfactory ways to derive the canonical distribution[3]):

$$\rho(0) = e^{-\zeta S^z}/\mathrm{Tr}\, e^{-\zeta S^z}, \tag{12}$$

where ζ is determined by the constraint (11). Incidentally, we remark that (12) could also be justified in a more traditional (but approximate) way, following more closely the actual preparation of the system in the laboratory as described in the first section (see again Ref. 1).

The problem now is to calculate

$$\langle S^z\rangle(t) = \mathrm{Tr}\, S^z \rho(t) \tag{13}$$

with

$$\rho(t) = U^t \rho(0) U^{-t}, \tag{14}$$

where $\rho(0)$ and U^t are prescribed by (12), (4), and (5). The expression (13) is more easily calculated if one writes it in the form

$$\langle S^z\rangle(t) = \mathrm{Tr}\, \{U^{-t} S^z U^t \rho(0)\}. \tag{15}$$

Since Tr is independent of the basis in which it is evaluated, we choose as a convenient basis:

$$\Psi_{\{\alpha_i\}} = \prod_i \otimes \psi_{\alpha_i}, \tag{16}$$

---

[2] G. Emch, Lecture notes, 8th Theoretical Physics Institute, University of Colorado, Summer 1965. Preprint *JILA*, University of Colorado and Natl. Bur. Std., Boulder, 1965. For further details see G. Emch, Helv. Phys. Acta **37**, 270 (1964); *ibid.*, **37**, 532 (1964); *ibid.*, **38**, 164 (1965), and references quoted therein. See also: G. Emch and C. Favre, Preprint Geneva 1965.

[3] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, 1955). See also the book published under the same title by G. W. Mackey (W. A. Benjamin Company, Inc., New York, 1963).

where $\{\psi_{\alpha_i}\}$ (with $\alpha_i = \pm 1$) is the basis defined in each $\mathfrak{H}^i$ by

$$\sigma^z \psi_{\alpha_i} = \alpha_i \psi_{\alpha_i}. \tag{17}$$

We next note that $\rho(0)$ is diagonal in the basis $\{\Psi_{\{\alpha_i\}}\}$, so that only the diagonal part of $(U^{-t}S^zU^t)$ is relevant for the evaluation of (15). Now, for any bounded operator A acting on $\mathfrak{H}$, it is convenient to define

$$\mathfrak{U}_n^t A \equiv U_n^{-t} A U_n^t. \tag{18}$$

Since the U_n^t commute among themselves,

$$\begin{aligned} U^{-t}S^zU^t &= \{\prod_{n\geq 0} \mathfrak{U}_n^t\} S^z \\ &= \sum_i \{\prod_{n\geq 0} \mathfrak{U}_n^t\} \sigma_i^z \\ &\equiv \mathfrak{U}^t S^z. \end{aligned}$$

Let us then first calculate

$$\mathfrak{U}_0^t \sigma_i^z = \{\prod_j \mathfrak{U}_{0,j}^t\} \sigma_i^z, \tag{19}$$

where

$$\mathfrak{U}_{0,j}^t A \equiv U_{0,j}^{-t} A U_{0,j}^t \tag{20}$$

with

$$U_{0,j}^t = \exp(+iB\sigma_j^z t). \tag{21}$$

Since

$$\sigma_j^z \text{ commutes with } \sigma_i^z \text{ for } j \neq i, \tag{22}$$

one has

$$\mathfrak{U}_0^t \sigma_i^z = \mathfrak{U}_{0,i}^t \sigma_i^z. \tag{23}$$

One can now use the property

$$(\sigma_i^z)^2 = I$$

to write

$$U_{0,i}^t = I \cos Bt + i\sigma_i^z \sin Bt. \tag{24}$$

Combining now (23) and (24) one gets:

$$\mathfrak{U}_0^t \sigma_i^x = \sigma_i^x \cos 2Bt + \sigma_i^y \sin 2Bt \tag{25}$$

and therefore

$$\begin{aligned} U^{-t}S^zU^t = (\{\prod_{n>0} \mathfrak{U}_n^t\} S^z) \cos 2Bt \\ + (\{\prod_{n>0} \mathfrak{U}_n^t\} S^y) \sin 2Bt. \end{aligned} \tag{26}$$

This takes care of the influence of the magnetic field. Next, we must calculate the influence of the spin–spin interaction. To do so, let us first evaluate

$$\mathfrak{U}_1^t \sigma_i^z = \{\prod_j \mathfrak{U}_{1,j}^t\} \sigma_i^z, \tag{27}$$

where $\mathfrak{U}_{1,j}^t$ is constructed in the usual way from

$$U_{1,j}^t = \exp[-i\epsilon(1)\sigma_j^z \sigma_{j+1}^z t]. \tag{28}$$

For reasons quite analogous to those encountered above, (27) reduces to

$$\mathfrak{U}_1^t \sigma_i^z = \mathfrak{U}_{1,i-1}^t \mathfrak{U}_{1,i}^t \sigma_i^z, \tag{29}$$

which can be readily calculated (using the same technique as for the B dependence) as:

$$\begin{aligned} \mathfrak{U}_1^t \sigma_i^z = \sigma_i^z \cos^2[2\epsilon(1)t] \\ - \tfrac{1}{2}\sigma_i^y(\sigma_{i-1}^z + \sigma_{i+1}^z) \sin[4\epsilon(1)t] \\ - \sigma_{i-1}^z \sigma_i^x \sigma_{i+1}^z \sin^2[2\epsilon(1)t]. \end{aligned} \tag{30}$$

At this point it becomes more and more intricate to write the successive explicit expressions for

$$\mathfrak{U}_2^t \mathfrak{U}_1^t \sigma_i^z, \qquad \mathfrak{U}_3^t \mathfrak{U}_2^t \mathfrak{U}_1^t \sigma_i^z, \qquad \text{etc.}$$

However, we should remember that we are not interested in the evaluation of the operator (26) for itself, but in the expectation value (15), to which only the diagonal part of (26) contributes. A somewhat closer glimpse to the form of the successive $\mathfrak{U}_n^t$ shows that the only part of

$$U^{-t}\sigma_i^z U^t$$

which contributes to (15) is

$$\sigma_i^z \prod_n \cos^2 2\epsilon(n)t.$$

The reader can rapidly convince himself that

$$U^{-t}\sigma_i^y U^t$$

cannot contribute to (15). We immediately have the desired result:

$$\langle S^z\rangle(t) = \langle S^z\rangle(0)[\prod_{n>0} \cos^2 2\epsilon(n)t] \cos 2Bt. \tag{31}$$

One also obtains

$$\langle S^y\rangle(t) = -\langle S^x\rangle(0)[\prod_{n>0} \cos^2 2\epsilon(n)t] \sin 2Bt, \tag{32}$$

$$\langle S^z\rangle(t) = \langle S^z\rangle(0) = 0. \tag{33}$$

The generalization of the above result from a one-dimensional chain to an n-dimensional crystal is obvious and can be taken care of simply by replacing $\epsilon(n)$ by ϵ_{ik} and making the subsequent trivial changes. The resulting expression is known and provides a qualitative (if not quantitative) agreement with experiment[1] if the ϵ_{ik} are properly adjusted. We gave the calculation with some details for the case of a linear chain with the following reasons:

(1) To emphasize that no approximation and no supplementary "statistical" assumptions are involved in the derivation of (31)–(33) when one starts from

(a) the generalized Ising-model microscopic evolution as prescribed by (4) and (5);

(b) the initial condition (12).

(This remains true also for any n-dimensional Ising-model generalized in the sense described above.)

(2) To provide a basis (in a hopefully intuitive language) for the generalized considerations, which is the topic of the following section.

III RELEVANCE TO THE GENERAL THEORY OF NONEQUILIBRIUM

A. Particularization of the model

The main purpose of Sec. II was to exhibit the presence in Eqs. (31)–(32) of the function

$$f(t) = \prod_{n>0} \cos^2 2\epsilon(n)t. \tag{34}$$

In general, one could try to discuss the behavior of this function for any $\epsilon(n)$ satisfying the very weak conditions (6). However, our purpose is not to produce the oddest analytical time behavior one could imagine from a particular model, but rather to exploit the greater simplicity of the model as extensively as possible in order to gain some insight into the general theory and the pseudoparadoxes and difficulties usually met. The very problem in which we are interested is to see how and why a macroscopic approach to equilibrium is compatible with a purely quantum mechanical microscopic description, in order to bring to light the central role played by the lack of information involved in any classical measurement on a quantum system.

We therefore postulate $\epsilon(n)$ to have the form

$$\epsilon(n) = \epsilon_0 2^{-n}. \tag{35}$$

The advantage of this choice is that $f(t)$ has an exceedingly simple form which can be deduced from a formula due to Euler:

$$f(t) = [\sin(\epsilon_0 t)/(\epsilon_0 t)]^2. \tag{36}$$

This is a positive function which obviously leads (with an oscillatory approach) to stationary value of $\langle S \rangle$:

$$\lim_{t\to\infty} f(t) \text{ exists and is zero!} \tag{37}$$

This result corresponds to an infinite linear chain with the spin-spin interaction extending over *all* pairs of neighbors. In the case of a finite ring, or equivalently, of an infinite chain with a cut-off described by

$$\epsilon(n) = 0 \quad \text{for all} \quad n > N, \tag{38}$$

one can also calculate exactly the corresponding

$$f_N(t) = \prod_{n=1}^{N} \cos^2 2\epsilon(n)t = f(t)W_N^{-1}(t), \tag{39}$$

where

$$W_N(t) = [\sin \epsilon(N)t/\epsilon(N)t]^2. \tag{40}$$

The function $W_N^{-1}(t)$ therefore, takes care of the "finite-size effects" (for the finite rings) or for the finite-extension of the interaction (for both finite rings and infinite linear chain). It shows that, for times $t \ll T_N$ with

$$T_N \approx 2^N \pi/\epsilon_0, \tag{41}$$

these effects are negligible. This remark is relevant for the recurrence problems and is confirmed by the fact that for finite times

$$\lim_{N\to\infty} W_N(t) = 1. \tag{42}$$

This, moreover, shows explicitly that, in the infinite-"volume" limit, the recurrence paradox resolves itself naturally without any recourse to more or less ill-defined probabilistic statements such as "only small deviations with the tendency of approaching equilibrium occur often."

B. Generalization of the model

One of the reasons why we explicitly did the calculations of Sec. II was to make obvious the fact that the model, as presented there, is open to generalizations. The first point we wish to make is that the only property we used of the initial state $\rho(0)$, as defined by (12), was the fact that $\rho(0)$ is diagonal in $\{\Psi_{\{\alpha i\}}\}$-basis (16). Therefore, the result (31)–(33) will remain true without any modification for quite a large class of initial states, namely for the $\rho(0)$ which are diagonal in (16).

Secondly, these results would also remain true if, instead of the one observable S^z, we would have been interested in any observable A of the form

$$A = \sum_i a_i \sigma_i^z. \tag{43}$$

Besides S^z, which is obviously recognized as a macroscopic observable, there is, therefore, a wide class of observables which also exhibit an approach to equilibrium, and are indeed related to a much finer

description of the system than the original one provided by S^z alone. In fact, the set generated by all of the observables of the form (43) is maximal Abelian. In a less pedantic language, one could simply express this property by saying that the set of the observables (43) is a (over) complete set of commuting observables (csco). More simply, the simultaneous measurement of all the observables which generate this set would lead to an information which cannot be improved by any compatible quantum measurement performed at that same instant. Using now a language even more familiar to statistical mechanicists, we could say that this generalized version of our model leads to an approach to equilibrium which does not require any real coarse-graining, or expressed in a better form, that the macrocell defined by the observation[2] are all one-dimensional. This corresponds to the situation usually referred to as *fine-graining*. That a definite approach to equilibrium is compatible with fine-graining illustrates a remark already made by Pauli[4] a long time ago. He emphasized that one of the few fundamental differences between classical and quantum statistical mechanics is that an information which is complete at a given instant remains so in time in the former description, whereas it can be lost in the latter. This is essentially due to the fact that, when the Hamiltonian does not commute with the csco considered, a state ρ, initially diagonal in the proper basis of this csco, does not remain diagonal; however, the nondiagonal elements of ρ which appears in the course of the evolution are of no relevance for the determination of the instantaneous expectation values of the observables belonging to the csco of interest. The generalized version of the model discussed in this subsection precisely provides an illustration of this remark. Incidentally, this is related to the fact that the quantum master equations, either fine- or coarse-grained, are formally identical.[2] We should also recall at this point that no random phase assumption is needed in the fine-grained case. This is consistent with the fact that the coarse-graining projector $\mathfrak{D}$, introduced in Ref. 2, coincides in this case (and in this case only) with the projector introduced by von Neumann[3] in his discussion of the measuring process.

C. Illustration of the coarse-graining concepts

We now consider the particular case of a (large but) finite ring, comprising say M sites whose positions are denoted by the index $i = 0, \cdots, M-1$. Each site is occupied by a $\frac{1}{2}$-spin particle. The time evolution of this system is again assumed to be given by (4) and (5) with the supplementary condition (35), complemented by any cut-off compatible with the ring structure. We moreover assume for simplicity that $B = 0$ (at least for $t > 0$). Now, instead of being interested in S^z we consider the following particular observable of the family (43)

$$A = \sigma_0^z = \sigma^z \otimes I \otimes \cdots \otimes I, \tag{44}$$

where σ^z is the usual two-dimensional Pauli matrix. The spectral decomposition of A is thus

$$A = \sum_{\Delta} A(\Delta) E_{\Delta} \tag{45}$$

with Δ running over the two indices $+$ and $-$. One has obviously

$$\begin{aligned} A(\pm) &= \pm 1, \\ E_{\pm} &= \tfrac{1}{2}(I \pm A), \end{aligned} \tag{46}$$

where $E_{\pm}$ are two orthogonal projectors with the following properties:

$$E_{\pm} \mathfrak{H}^i \subseteq \mathfrak{H}^i \tag{47}$$

(the equality sign is valid for all i except $i = 0$) and

$$\dim E_{\pm} = 2^{M-1}.$$

One can now use the language systematically established in Ref. 2 and say that the $E_{\pm}$ are the two *macrocells* of our system: An observation based on A only can provide an information on the relative populations of these macrocells, but cannot lead to any information on the inside of them. One can furthermore introduce the maximal representative of the two macroscopic equivalence classes which occur in this model:

$$W_{\Delta} = E_{\Delta}/2^{M-1}. \tag{48}$$

The macroscopic state $\{p_{\Delta}^t\}$ corresponding to any given microscopic state W^t is defined by:

$$p_{\Delta}^t = \mathrm{Tr}\, W^t E_{\Delta}. \tag{49}$$

Following the considerations developed in Ref. 2 we are only interested (for the prediction of the evolution of the expectations values of A) in the time-dependent p_{Δ}^t with the initial condition

$$W^0 = \sum_{\Delta} p_{\Delta}^0 W_{\Delta} \tag{50}$$

and the microscopic evolution equation

$$W^t = U^{-t} W^0 U^t.$$

[4] W. Pauli, Nuovo Cimento, Suppl. 6, 166 (1949).

This problem can be solved directly (without the need to go through the master equation techniques) following the calculations of Sec. II. We obtain

$$p_{\pm}^{t} = \tfrac{1}{2}(1 \pm \alpha f_N(t)), \tag{51}$$

where α is determined by

$$p_{\pm}^{0} = \tfrac{1}{2}(1 \pm \alpha). \tag{52}$$

In this form the model will lead to further considerations.

D. Ergodicity

The first problem of interest in the discusssion of nonequilibrium problems is to check whether a given system approaches equilibrium in some very restricted sense, namely

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T dt\, \langle A\rangle(t) = \overline{\langle A\rangle}, \tag{53}$$

where $\overline{\langle A\rangle}$ is the *microcanonical* equilibrium value of the observable A. Although it is clear that the left-hand side of Eq. (53) has hardly any physical (operational) sense in connection with the problem of the true approach to equilibrium, its evaluation has some relevance to the problem, for the following reason: if $\langle A\rangle(t)$ ever approaches a limit as $t \to \infty$, then this limit should be equal to the ergodic average of $\langle A\rangle(t)$ which by definition is the left-hand side of (53) [There is no point to discuss here ergodic limits other than the ordinary $(C, 1)$—Cesaro average.]

One could calculate the left-hand side of (53) directly from (51). One would then see immediately that (53) is satisfied when the length of the cut-off (described by N) and subsequently the size of the ring (described by M) both go to infinity

$$\lim_{t\to\infty} \lim_{N\to\infty} f_N(t) = \lim_{t\to\infty} f(t) = 0, \tag{54}$$

and consequently

$$\lim_{t\to\infty} \lim_{N\to\infty} p_{\pm}(t) = \tfrac{1}{2}, \tag{55}$$

which corresponds to the microcanonical distribution.

We now discuss another approach which can be illustrated by the model in the form discussed in Sec. III.C. A general criterion for (53) has been proposed[2] in the case of discrete spectrum of the Hamiltonian. This criterion reads as follows [see also (68) below]:

$$\sum_{\mu} \mathrm{Tr}\,(P_\mu W_\Delta P_\mu E_{\Delta'}) = N_{\Delta'}/N_3 \quad \text{for all} \quad E_\Delta, E_{\Delta'} \subset S_3, \tag{56}$$

where W_Δ is defined as above [by (48) in the present case], $N_\Delta = \dim E_\Delta$, N_3 is the dimension of the energy-shell S_3, and P_μ are the eigenprojectors of the total Hamiltonian.

To make this condition clear, we first have to state precisely what we mean by energy-shell. An energy-shell is defined as any eigenprojector of the following operator

$$H_0 \equiv \sum_{\Delta} (\mathrm{Tr}\, H W_\Delta) E_\Delta, \tag{57}$$

to which we refer as the *macroscopic energy* for reasons which have been explained.[2]

In the present model

$$H_0 = 0, \tag{58}$$

and therefore there is only *one* energy-shell: the whole space $\mathfrak{H}$ defined by (3). We then have

$$N_3 = 2^M. \tag{59}$$

To see whether (56) is satisfied or not, we proceed as follows: First, we remark that the whole Hamiltonian of the system as it results from the definition (4), (5) is not relevant for the evolution of our model, but can be replaced by the effective Hamiltonian

$$H_e = \sigma_0^z \sum_{n>0} \epsilon(n)(\sigma_n^z + \sigma_{M-n}^z) \tag{60}$$

without changing anything to the evolution, as it should be clear from the detailed calculations of Sec. II. This Hamiltonian is diagonal in the basis of $\mathfrak{H}$ defined by

$$\Phi_{\{\beta_i\}} = \prod_i \otimes \varphi_{\beta_i}, \tag{61}$$

where

$$\sigma^z \varphi_{\beta_i} = \beta_i \varphi_{\beta_i} \tag{62}$$

with

$$\beta_i = \pm 1. \tag{63}$$

Let us now remark that

$$P_{\{\beta_i\}} W_\pm P_{\{\beta_i\}} = P_{\{\beta_i\}}/N_3 \quad \text{for all} \quad \{\beta_i\}, \tag{64}$$

where $P_{\{\beta_i\}}$ are the one-dimensional projectors on the respective pure states $\Phi_{\{\beta_i\}}$.

Therefore, if the spectrum of H_e were nondegenerate, (56) and consequently (53) would be satisfied. However, in the model considered here, H_e is degenerate and, consequently, there appear terms of the form

$$P_{\{\beta_i\}} W_\pm P_{\{\beta'_i\}} \tag{65}$$

in the summation (56), where $\{\beta_i\}$ and $\{\beta_i'\}$ would correspond to configurations having the same total energy.

Given any configuration $\{\beta_i\}$, let us denote by $\{\beta_i\}^*$ the configuration obtained from $\{\beta_i\}$ by changing only β_0 into $(-\beta_0)$. Let us, furthermore, denote by p_{β_i} (resp. w_Δ) the restriction of the operator $P_{\{\beta_i\}}$ (resp. W_Δ) to the subspace $\mathfrak{H}^t$ (resp. $\mathfrak{H}^0$). We can now remark that (65) vanishes unless

$$\{\beta_i'\} = \{\beta_i\}^* \tag{66}$$

and that (65) is then equal to

$$(p_{\beta_0} w_\Delta p_{\beta_0}) \otimes (\prod_{i \neq 0} \otimes p_{\beta_i}). \tag{67}$$

From the symmetry of the model, it is, furthermore, evident that two configurations $\{\beta_i\}$ and $\{\beta_i\}^*$ can correspond to the same eigenvalue of H_e only if this eigenvalue is zero. Therefore, for all eigenprojectors P_μ corresponding to a nonzero eigenvalue of the effective Hamiltonian, we have

$$P_\mu W_\Delta P_\mu = P_\mu / N_\Im \tag{68}$$

valid for all P_μ except P_0 (the eigenprojector corresponding to the eigenvalue zero of H_e). At this point, we draw the attention of the reader to the importance of Eq. (68) (satisfied for *all* P_μ!) in the general theory developed in Ref. 2 for the existence and the uniqueness of an equilibrium state. The fact that P_0 does not satisfy (68) allows us, then, to evaluate the departure from ergodicity (in the sense of Ref. 2) in our model. It is, in fact, a simple matter of playing with the degenerate eigenstates corresponding to the eigenvalue zero to evaluate the contribution of this level to the ergodic average. One finally obtains

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T dt\, p_\pm^t = \tfrac{1}{2}[1 \pm (\tfrac{1}{2})^N (p_+^0 - p_-^0)]. \tag{69}$$

This result completes (55), which was obtained more directly. We remark that, besides its theoretical interest, this evaluation of the ergodic average in the case of a finite ring (or of an infinite linear chain with a cut-off in the interaction) can be generalized quite easily to the case where $\epsilon(n)$ has a more general form than that assumed in (35). This allows us, in particular, to determine, without calculating the exact time dependence of the expectation value of the magnetization, the value around which it oscillates and which is usually referred to as its equilibrium value.[2] In the case of an infinite chain with an interaction extending to infinity, this equilibrium value is zero, the magnetization then oscillates above its equilibrium value and, in fact, approaches it with a damped oscillating time behavior. That this equilibrium value corresponds to the microcanonical distribution is readily seen.

In the above remarks we were careful to describe the approach to equilibrium in terms of expectation values of macroscopic observables. It is not by accident that we avoided statements in terms of microscopic density operators. The reason for this is that it is only in the macroscopic sense that the states considered here approach equilibrium. More precisely, given any state W^0 of the form (50), we do not assert that W^t approaches, in general, the microcanonical distribution in the course of time. We only claim that the macroscopic equivalence class of W^t [defined through (49) as the set $\{p_\Delta^t\}$] approaches the macroscopic equivalence class (defined by $p_\Delta = N_\Delta/N_\Im$) of the microcanonical distribution. This statement is obviously sufficient to account for the *macroscopic approach to microcanonical equilibrium*, even if no such approach can be traced when one uses only arguments pertaining to the microscopic description. By these considerations we hope to strongly emphasize the central role of the concept of macroscopic equivalence classes of microscopic states inherited from the notion of coarse-graining (even in cases where not all the assumptions usually made[2] are satisfied!).

E. Zermelo's and Loschmidt's Paradoxes

We already mentioned in Sec. III.A how this model allows us, without any recourse to probabilistic arguments, to eliminate the recurrence paradox by a passage to the limit of infinite size and infinite extension of the interaction, and then obtain a bona fide approach to equilibrium. This takes care of Zermelo's paradox in the sense predicted by Boltzmann (and not only in a probabilistic sense!) in his well-known exclamation "Then, wait!" The expression (40) states precisely "how long!"

This model also shows how an approach to equilibrium can be compatible with the reversibility of the microscopic evolution.

This reversibility manifests itself through

$$f(t) = f(-t) \tag{70}$$

and, incidentally, also through

$$f_N(t) = f_N(-t) \tag{71}$$

(which, however, is less surprising). Equation (70) can be interpreted in two ways. One could say that the behavior for $t < 0$ is simply irrelevant since it cannot be produced in the laboratory. This is,

however, not quite satisfactory and we feel that a better interpretation seems to be the following: Given a situation at time $t = 0$, we have the same information on "which situation it comes from" as we have on "which situation it will evolve to." In short, the security of a prediction is the same as that of a postdiction. This illustrates also the challenge raised when Boltzmann reportedly answered to the critics of Loschmidt's type: "Go ahead, reverse them." Our model indicates a slightly more sophisticated answer which we believe to be of quite a general character: If we "reverse the time" at a given instant, say $t = t_0$, taking only into account the macroscopic knowledge we have of the system at *this particular instant* $t = t_0$, we shall also have an approach to equilibrium in the direction of the past, and the system will in general not pass again through its initial state (the verification of this last statement cannot be made since it was decided, at $t = t_0$, that any information besides the one at that time $t = t_0$ is forgotten). On the contrary, if we keep records of what happened on the system between $t = 0$ and $t = t_0$, and then reverse the time at $t = t_0$, taking into account all the information we now have, the system will return to its initial state (at time $t = 0$) and then decay to equilibrium in the direction of the past in agreement with (70).

F. The Non-Markovian Character of the Evolution

In statistical mechanics, one of the favorite ways to obtain (with relative ease) an equation describing (at least within some approximation) the actual approach to equilibrium is to make an assumption of the kind of Boltzmann's *Stosszahlansatz*. The quantum analog of this assumption is the so-called repeated random phase assumption. Pauli used it to derive his famous master equation, hereafter referred to as the PEM. The nonmechanical character of this assumption is evident. Its result is that the reduction of the microscopic unitary evolution to the macroscopic subspace[2] becomes a semigroup, the generator of which is dissipative in case of an actual approach to equilibrium. The evolution obtained in this way is usually referred to as a Markovian process. The validity of this assumption has been extensively questioned in the last ten years.

The main property of the present model, in the different variations presented here, is that no repeated random phase is required to solve it exactly and to obtain an actual approach to equilibrium. The fact that such an assumption is not needed in the derivation of the result still does not prove that the evolution is non-Markovian. To show this we proceed *ad absurdo*. Suppose then that the evolution were Markovian. Under the usual continuity assumptions (which are satisfied here!) the evolution of a Markov process is determined by the Chapman–Kolmogorov equations. These equations are nothing but the mathematical abstraction corresponding to the PEM. In a perhaps not quite orthodox[2] form (but equivalent to the usual one), the PEM can be written as

$$(d/dt)p^t = -\Lambda p^t \tag{72}$$

defined only for $t \geq 0$. Λ is an operator acting in the *macroscopic* Liouville subspace.[2] Λ is positive and Hermitian in the usual PEM. When Λ is bounded, (72) can be integrated (without any "if" and "but") to give

$$p^t = e^{-\Lambda t}p^0. \tag{73}$$

If, moreover, Λ has a discrete spectrum, the p_Δ^t are discrete superpositions of nonincreasing real exponentials.

In the variation of our model considered in Sec. III.C, the macroscopic space is two-dimensional. Therefore Λ is obviously bounded and has at most two distinct eigenvalues. As a consequence, the p_Δ^t, if described by a PEM, should be the superposition of at most two nonincreasing real exponentials. We, however, have the explicit form (51) of the p_Δ^t. Even in case where both the size (M) of the system and the extension (N) of the interaction are infinite, these p_Δ^t cannot be written as linear combination of two nondecreasing real exponentials. Consequently, by this reduction *ad absurdo*, we proved that this variation of the model presents a macroscopic approach to equilibrium which cannot be described as a Markovian process. This result apparently depends on the form of the function $\epsilon(n)$ as ascribed by (35). It is, in fact, the expression of a general theorem.[2] One could now still argue further. Very often one sees that a system, the evolution of which has to be described by a generalized master equation (GEM), can, however, present an evolution which, in the long-time limit, can be described by a PEM, at least if one does some kind of time-smoothing.[2] This, however, is not the case here, as might readily be seen from the time-smoothed, long-time limit of (51), even if one first carries out the limit of infinite extension of the interaction. This last result now depends *strongly* on the special assumption (35). In particular, it can be seen that there exist some other particular spatial dependence of the interaction (which are indeed closer to the actual spatial dependence of the dipole-dipole in-

teraction!) which would lead to an asymptotically exponential decay to equilibrium.

From the point of view of the axiomatics, the interest of this version of the model, as discussed above, is that it exhibits explicitly an approach to equilibrium which can be predicted exactly, and which can in *no sense* be described (even in some approximation) as a Markov process, the differential evolution equation of which is of the PEM type. Therefore, if one wants to describe the evolution of the model through a master equation, one has to consider a GEM which cannot be approximated by a PEM. One could obviously carry out the evaluation of the kernel of the GEM for this model. This would, however, turn out to be a somewhat tedious task, especially in the case of an interaction extending in space as (35). The simplicity of the model would anyway be lost in the process as one could already figure out by feeding the result (51) into the Laplace transform of the GEM. Similarly, one could also discuss on this model van Hove's conditions of diagonal singularity and interconnection of states. This would, however, not lead to an information deeper than that obtained from the general theory.[2] We therefore do not pursue further in this direction here.

IV CONCLUSIONS

The formal simplicity of a generalized Ising model allowed us to calculate exactly the time evolution of the transverse magnetization as well as some other related quantities, starting from a wide class of initial conditions. The exact solubility of the model was used to discuss several problems connected with the actual approach to equilibrium. We succeeded in stating precisely some of the persistently unclear statements related to the phenomena.

The two main ingredients of nonequilibrium statistical mechanics, namely the unitary, mechanistic, microscopic evolution and the partial macroscopic information connected already with any classical measurement on a quantum system, were proved to be sufficient (in principle) to ensure in some cases an actual approach to equilibrium if one evaluates correctly the infinite-size limit. In particular, the nonnecessity of any kind of repeated random phase assumption was exemplified by a truly non-Markovian model.

ACKNOWLEDGMENTS

The author wishes to express his gratitude to Professor W. E. Brittin for the hospitality extended to him as a Visiting Fellow of the Joint Institute for Laboratory Astrophysics of the University of Colorado and National Bureau of Standards at Boulder, where this work was undertaken. He also wishes to acknowledge the discussions he had during this period with Dr. E. G. D. Cohen, Dr. R. L. Peterson, and Dr. R. J. Swenson.[5]

Thanks are also due to Dr. E. W. Montroll and Dr. N. S. Goel for their comments as well as for their reading of the manuscript.

---

[5] When this work was completed, Dr. R. J. Swenson drew the author's attention to the following paper: S. P. Heims, Am. J. Phys. **33**, 722 (1965), the content of which overlaps with the first two sections of the present report. These two sections have been maintained in this publication because of the fact that the preliminary calculations are presented here in a way to provide a guide for the subsequent considerations.