An extension of the Kac ring model

2003 J. Phys. A: Math. Gen. 36 11547

(http://iopscience.iop.org/0305-4470/36/46/001)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **36** (2003) 11547–11559 PII: S0305-4470(03)66392-4

An extension of the Kac ring model

Wojciech De Roeck[1,3], **Tim Jacobs**[1], **Christian Maes**[1] **and Karel Netočný**[2]

[1] Instituut voor Theoretische Fysica, K U Leuven, B-3001 Leuven, Belgium
[2] Instituut voor Theoretische Natuurkunde, Rijksuniversiteit Groningen, the Netherlands

E-mail: tim.jacobs@fys.kuleuven.ac.be

Received 18 July 2003
Published 5 November 2003
Online at stacks.iop.org/JPhysA/36/11547

Abstract
We introduce a unitary dynamics for quantum spins which is an extension of a model introduced by Mark Kac to clarify the phenomenon of relaxation to equilibrium. When the number of spins becomes very large, the magnetization satisfies an autonomous equation as a function of time with exponentially fast relaxation to the equilibrium magnetization as determined by the microcanonical ensemble. This is proved as a law of large numbers with respect to a class of initial data. The corresponding Gibbs–von Neumann entropy is also computed and its monotonicity in time discussed.

PACS numbers: 05.30.Ch, 05.40.−a

1. Relaxation to equilibrium

The Kac ring model was introduced by Mark Kac to clarify how manifestly irreversible behaviour can be obtained from an underlying reversible dynamics [2, 6, 9]. It explains via a simple model some of the conceptual subtleties in the problem of relaxation to equilibrium such as, for example, are present in the derivation and the status of the Boltzmann equation for dilute gases. In particular, the Kac dynamics shares some basic features with a Hamiltonian time evolution such as being deterministic and dynamically reversible. In this paper we extend this dynamics to a unitary evolution on a finite quantum spin system. Again, the dynamics remains far from realistic but it allows a precise formulation and discussion of some features of relaxation to equilibrium for a quantum dynamics. This is especially useful and relevant as, in the quantum domain, the problem of relaxation is beset with even greater conceptual difficulties. In our framework, relaxation to equilibrium becomes visible if one can select a small number of macroscopic variables that typically evolve via autonomous deterministic equations to take on values that correspond to equilibrium. Typical refers to a law of large numbers with respect to the initial data. Paradoxes are avoided by taking seriously the fact that relaxation is a macroscopic phenomenon, involving a huge number of degrees of freedom

[3] Aspirant FWO Vlaanderen, UI Antwerpen.

whose evolution is monitored over a realistic time span. One should also keep in mind that relaxation to equilibrium goes beyond questions of return to equilibrium, see [7], which are mostly related to stability of equilibrium states. A more general introduction to this and various related problems can be found in [8].

In section 2, we introduce the model and we state the basic result. Section 3 is devoted to a discussion of related issues. The proofs are postponed to section 4.

2. Model and results

2.1. The model

Consider N sites on a ring (periodic boundary conditions). Between any two neighbouring sites there is a fixed *scattering* mechanism to be specified below, and at each site, we find a spin-$\frac{1}{2}$ particle. Time is discrete and at each step the ring rotates in a fixed direction over one ring segment. Depending on the segment that each spin crosses, it is scattered to another state.

For the Hilbert space $\mathcal{H}_N$ we take the N-fold product of copies of $\mathbb{C}^2$, the state space at each site:

$$\mathcal{H}_N \equiv \bigotimes_{j=1}^{N} \mathbb{C}^2_j$$

with the standard inner product that defines the Hermitian conjugate for matrices denoted by the superscript $\star$. We write the elements of $\mathcal{H}_N$ as vectors in $\mathbb{C}^{2N}$, denoted by η or η' with components η_j indicating the state of the spin at site j.

2.1.1. Dynamics and observables. We fix an arbitrary Hermitian matrix $H \in \mathbb{C}^{2\times 2}$, called a single-site Hamiltonian, and construct the unitary matrix

$$U \equiv \mathrm{e}^{\mathrm{i}H} \qquad U^\star = \mathrm{e}^{-\mathrm{i}H}. \tag{2.1}$$

With each segment on the ring connecting two nearest neighbours $(j,\ j+1)$ there is associated a scattering variable $\epsilon_j \in \{0, 1\}$. The configuration of scatterers is denoted by $\epsilon \equiv (\epsilon_1, \ldots, \epsilon_N)$ and does not change in time. We define the unitary matrix $U_j^{\epsilon_j}$ on $\mathbb{C}^2_j$ as

$$U_j^{\epsilon_j} \equiv \begin{cases} \mathbb{1} & \text{for} \quad \epsilon_j = 0 \\ U & \text{for} \quad \epsilon_j = 1. \end{cases}$$

The superscript ϵ_j can thus be read as a power and $U_j^{\epsilon_j} = \exp(\mathrm{i}\epsilon_j H_j)$ with H_j a copy of H working on $\mathbb{C}^2_j$. This defines the scattering mechanism.

The rotation of the ring is implemented by the operation R on $\mathcal{H}_N$, which transforms every vector $\eta = (\eta_1, \eta_2, \ldots, \eta_N)$ into

$$R\eta \equiv (\eta_N, \eta_1, \ldots, \eta_{N-1}) \in \mathcal{H}_N$$

and $R^\star = R^{-1}$, rotation in the opposite direction. Combining this with the scattering mechanism finally gives rise to the unitary operator $U_N \equiv U_N(\epsilon)$ on $\mathcal{H}_N$ via

$$U_N \equiv R \bigotimes_{j=1}^{N} U_j^{\epsilon_j}.$$

To be specific we look at positive times and the state at time $t = 1, 2, \ldots$ is then obtained from the state $\eta_0 \in \mathcal{H}_N$ at time $t = 0$ from

$$\eta_t \equiv U_N(t)\eta_0 = U_N(t-1)U_N\eta_0.$$

This unitary dynamics on the finite-dimensional $\mathcal{H}_N$ is rather simple. For many observables (such as the total magnetization) the dynamics is entirely equivalent to leaving the spins in place and rotating the scatterers instead. We therefore work with the more convenient

$$U_N(t) = \bigotimes_{j=1}^{N} U_j^{k_j(\epsilon,t)} \tag{2.2}$$

with $k_j(\epsilon, t) \equiv \sum_{n=1}^{t} \epsilon_{j-n}$ (modulo N). The updating of the spins is independent modulo the fact that they may have a scattering mechanism in common.

The Hamiltonian is a sum of single-site contributions,

$$H_N \equiv \sum_{j=1}^{N} H_j \tag{2.3}$$

with $H_j \equiv \mathbb{1} \otimes \cdots \otimes H \otimes \cdots \otimes \mathbb{1}$ (on the jth site) and for concreteness, we decompose the single-site Hamiltonian in the Pauli basis

$$H = h_1\sigma_x + h_2\sigma_y + h_3\sigma_z = \vec{h} \cdot \vec{\sigma} \qquad 0 \neq \vec{h} \in \mathbb{R}^3. \tag{2.4}$$

The eigenvalues are $e_+ = \|\vec{h}\| \equiv h$ and $e_- = -e_+$. (We can of course ignore adding a constant to this Hamiltonian.) Our observables are also built up by adding one-site contributions. We are given a Hermitian matrix $A \in \mathbb{C}^{2\times 2}$ and we construct the system operators

$$A_N \equiv \sum_{j=1}^{N} \mathbb{1} \otimes \cdots \otimes A \otimes \cdots \otimes \mathbb{1} \tag{2.5}$$

where A is at the jth position. We will consider the magnetization vector $\vec{M}_N \equiv \left(M_N^x, M_N^y, M_N^z\right)$ defined through the one-site observables $M^\alpha = \sigma_\alpha$, the Pauli matrices, with $\alpha \in \{x, y, z\}$.

The restriction of looking only at one-site observables will be discussed in section 3.4.

2.1.2. Equilibrium. The dynamics depends non-trivially on the scatterers but the equilibrium properties are independent of them. This is similar to the situation in the Boltzmann–Grad limit where the hard core matters dynamically but does not enter in the computations of energy or pressure. Energy is conserved in the sense that the Hamiltonian commutes with the time evolution: $[U_N(t), H_N] = 0$.

Equilibrium is characterized by the microcanonical distribution. Let ψ_+, ψ_- be the two eigenvectors of H with eigenvalues $e_+ \geqslant e_-$ and spectral projectors P_{e_+} and P_{e_-}, respectively. For arbitrary $e \in [e_-, e_+]$ we select an energy space by

$$P_e^N \equiv \frac{1}{Z_e^N} \sum_{e} P_{e_1} \otimes \cdots \otimes P_{e_N} \tag{2.6}$$

where, in $\sum_e$, we sum over all (e_j) with fixed $\sum_j e_j$ satisfying $Ne \leqslant \sum_j e_j < e_+ - e_- + Ne$; the normalization Z_e^N ensures that $\mathrm{Tr}\left[P_e^N\right] = 1$. Alternatively, we could have summed in (2.6) over all (e_j) with $\sum_j e_j/N$ in a certain interval around e and at the very end let the interval shrink to zero. The average

$$\langle A_N \rangle_e^N \equiv \mathrm{Tr}\left[P_e^N A_N\right] \tag{2.7}$$

for a Hermitian matrix A_N on $\mathcal{H}_N$ defines the (finite volume) microcanonical ensemble.

The (infinite volume) equilibrium magnetization corresponding to an energy e is then defined from (2.7) to be the limit

$$\vec{m}_e \equiv \lim_{N\uparrow+\infty} \frac{1}{N} \langle \vec{M}_N \rangle_e^N. \tag{2.8}$$

This limit exists and can easily be computed: it equals

$$\vec{m}_e = e\frac{\vec{h}}{h^2}. \tag{2.9}$$

2.1.3. Initial data. Initial data are determined by a density matrix ρ_0^N on $\mathcal{H}_N$ and by the choice of the scatterers ϵ. Concerning the state of the spins, let us keep in mind an initial preparation with a particular magnetization in the z-direction, not in equilibrium. Let Q_+ and Q_- be the projectors on spin up and spin down, respectively, with $\sigma_z = Q_+ - Q_-$, to define

$$Q_m^N \equiv \sum_m Q_{m_1} \otimes \cdots \otimes Q_{m_N} \qquad m \in \{-1, +1\} \tag{2.10}$$

where, as in (2.6), in $\sum_m$, we sum over all (m_j) with fixed $\sum_j m_j$ satisfying $Nm \leqslant \sum_j m_j < 2 + Nm$. Equation (2.10) gives the projector on the magnetization space with magnetization converging to m as $N \uparrow +\infty$. A possible initial density matrix is then ($0 < d(m, N) < +\infty$ is a normalization)

$$\rho_0^N = \frac{1}{d(m, N)} Q_m^N. \tag{2.11}$$

Another (but thermodynamically equivalent) choice would be the grand-canonical

$$\rho_0^N = \frac{1}{Z_N(\lambda)} \exp\left(\lambda M_N^z\right) \tag{2.12}$$

and similarly for the other directions of magnetization. A more general class of initial density matrices ρ_0^N will be introduced at the beginning of section 2.2. Most important is that they satisfy a law of large numbers for observables of the form (2.5) and that the single-site marginal remains well defined in the thermodynamic limit, see (2.22).

Now the dynamics starts and the density matrix at time t is defined from (2.2):

$$\rho_t^N \equiv \rho_t^N(\epsilon) \equiv U_N(t)\rho_0^N U_N(t)^\star \tag{2.13}$$

and now also depends on the scatterers ϵ_j.

We are interested in the magnetization in ρ_t^N and in its limiting behaviour as (first) $N \uparrow +\infty$ and (then) $t \uparrow +\infty$. In order to take these limits, we also need to specify the initial (and unchanging) condition of the scatterers. For this we introduce the set

$$\Omega_\mu^N \equiv \left\{ \epsilon \in \{0, 1\}^N \,\middle|\, \sum_{j=1}^N \epsilon_j = \lceil \mu N \rceil \right\} \tag{2.14}$$

with $\lceil \mu N \rceil$ the integer part of $\mu N + 1$. The constraint fixes the fraction of *active* scatterers to be about μ. The probability of selecting one particular ϵ in Ω_μ^N is uniform:

$$\mathbb{P}_\mu^N(\epsilon) \equiv \frac{1}{\binom{N}{\lceil N\mu \rceil}}. \tag{2.15}$$

All limits will involve this probability distribution. We say that a sequence of functions G_N on Ω_μ^N typically takes the value g_μ, written as

$$\epsilon\text{-}\lim_{N\uparrow+\infty} G_N = g_\mu \qquad \text{iff} \quad \lim_{N\uparrow+\infty} \mathbb{E}_\mu^N[(G_N - g_\mu)^p] = 0 \qquad p = 1, 2 \tag{2.16}$$

where the expectation refers to the probabilities (2.15). Via the Chebyshev inequality one can reformulate (2.16) as a statement about the fraction of scatterers in Ω_μ^N for which $G_N \simeq g_\mu$. This relation is made stronger if higher p in (2.16) are obtained. For simplicity, we restrict ourselves here to the average ($p = 1$) and to the variance ($p = 2$).

2.1.4. Entropy. One expects for the given model that the one-particle distribution satisfies an autonomous equation. This information is encoded in the single-site density matrix. For finite N, the marginals are defined through

$$\nu^N \equiv \frac{1}{N}\sum_{j=1}^{N} \mathrm{Tr}_j[\rho^N] \tag{2.17}$$

with Tr_j the reduced density matrix at site j. The Gibbs–von Neumann entropy is defined following the ideology of [3] through the variational principle

$$S_N(\rho^N) \equiv \sup_{\rho'^N} -\mathrm{Tr}[\rho'^N \log \rho'^N] \tag{2.18}$$

where the supremum is over all density matrices ρ'^N with marginal ν^N from (2.17). Obviously, the supremum in (2.18) is attained for the product state $\rho'^N = \bigotimes \nu^N$ and therefore through (2.17), the quantum entropy equals

$$S_N(\rho^N) = -N\,\mathrm{Tr}[\nu^N \log \nu^N]. \tag{2.19}$$

We insert ρ_t^N from (2.13) in (2.19) and put, hoping all goes well,

$$s(t) \equiv \epsilon\text{-}\lim_{N\uparrow+\infty} \frac{1}{N} S_N\left(\rho_t^N\right) \tag{2.20}$$

as the time-dependent entropy (density). The more frequently considered von Neumann entropy $-\mathrm{Tr}\left[\rho_t^N \log \rho_t^N\right]$ does not change in time and is therefore less relevant here.

2.2. Results

The conditions for our main result on the dynamics (2.13) are on the level of initial data, see section 2.1.3. We say that ρ_0^N is a density matrix that is dispersion-free in the sense that for all observables (2.5) and all continuous f

$$\lim_{N\uparrow+\infty} \left| \mathrm{Tr}\left[f\left(\frac{A_N}{N}\right)\rho_0^N\right] - f\left(\mathrm{Tr}\left[\frac{A_N}{N}\rho_0^N\right]\right)\right| = 0. \tag{2.21}$$

Furthermore, we need the existence of the thermodynamic limit of the system-averaged one-site marginal defined in (2.17):

$$\nu_0 \equiv \lim_{N\uparrow+\infty} \frac{1}{N}\sum_{j=1}^{N} \mathrm{Tr}_j\left[\rho_0^N\right] \equiv \frac{\mathbb{1} + \vec{m}_0\cdot\vec{\sigma}}{2} \tag{2.22}$$

and similarly, we also define the time-evolved version of ν_0, i.e.,

$$\nu_t \equiv \epsilon\text{-}\lim_{N\uparrow+\infty} \frac{1}{N}\sum_{j=1}^{N} \mathrm{Tr}_j\left[\rho_t^N\right]. \tag{2.23}$$

Finally, the sequence ρ_0^N must satisfy a technical condition. We write

$$\rho_0^N = \sum_{w=1}^{r(N)} \bigotimes_{j=1}^{N} \chi_{j,w}^N \tag{2.24}$$

the initial density matrix as a sum of products of 2×2 matrices, and we assume that

$$\sup_{N\in\mathbb{N}} \sup_{\substack{w=1,\ldots,r(N)\\ j=1,\ldots,N}} \left\|\chi_{j,w}^N\right\| \equiv C < \infty. \tag{2.25}$$

These three conditions are obviously satisfied by (2.12). That (2.21) is also satisfied for (2.11) needs an argument given in the example after the proof of lemma 4.3.

The initial magnetization is denoted by $\vec{m}_0 \equiv \mathrm{Tr}[\vec{\sigma}\,\nu_0]$ and the equilibrium magnetization $\vec{m}_e$ is defined in (2.8). We denote components of the magnetization as m^α, $\alpha \in \{x, y, z\}$. The initial energy e is found from $e \equiv \mathrm{Tr}[H\nu_0]$.

Theorem 2.1. *For all continuous functions f and for initial ρ_0^N satisfying the above conditions,*

$$\underset{N\uparrow\infty}{\epsilon\text{-}\lim}\,\mathrm{Tr}\left[f\left(\frac{M_N^\alpha}{N}\right)\rho_t^N\right] = f\left(m_t^\alpha\right) \tag{2.26}$$

with

$$\vec{m}_t \equiv \vec{m}_e + \mathrm{Re}\left[\left(\vec{m}_0 - \vec{m}_e + \frac{\mathrm{i}(\vec{h}\times\vec{m}_0)}{h}\right)(1-\mu+\mu\exp[\mathrm{i}(e_+ - e_-)])^t\right]. \tag{2.27}$$

In particular, for $f(x) = x$, the magnetization

$$\lim_{t\uparrow+\infty}\underset{N\uparrow+\infty}{\epsilon\text{-}\lim}\frac{1}{N}\mathrm{Tr}\left[\vec{M}_N\rho_t^N\right] = \vec{m}_e \tag{2.28}$$

converges exponentially fast to its equilibrium value, for all $\mu \in (0, 1)$, when $e_+ - e_-$ is not a multiple of 2π.

The entropy (2.20) equals

$$s(t) = -\frac{1+\|\vec{m}_t\|}{2}\log\left(\frac{1+\|\vec{m}_t\|}{2}\right) - \frac{1-\|\vec{m}_t\|}{2}\log\left(\frac{1-\|\vec{m}_t\|}{2}\right) \tag{2.29}$$

where $\vec{m}_t$ is given by (2.27). $s(t)$ increases with time (H-theorem) and can be written as

$$s(t) = \mathcal{S}[\nu_t] \qquad \mathcal{S}[\nu] \equiv -\mathrm{Tr}[\nu\log\nu]. \tag{2.30}$$

3. Discussion

3.1. Kac ring model

The Kac ring model in [6], p 99, can be recovered by the appropriate choice of the Hamiltonian (2.4), by only looking at the magnetization in the z-direction. Identifying $\eta_i = 1$ or $|\uparrow\rangle$ with a white ball and the state $\eta_i = -1$ or $|\downarrow\rangle$ with a black ball in the original ring model, the switching between a white and a black ball corresponds to a spin flip. This is accomplished by the single-site Hamiltonian (written in the σ_z base):

$$H = \frac{\pi}{2}\begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix} \quad\Rightarrow\quad U = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}$$

which, up to an irrelevant constant, corresponds to the choice $\vec{h} = -\left(\frac{\pi}{2}, 0, 0\right)$ in (2.4). The eigenvalues of this Hamiltonian satisfy $e_+ - e_- = \pi$ and the equilibrium magnetization is $m_e^z = 0$. We start by randomly sampling the balls (or spins) with a fixed overall colour (or magnetization) in (2.11) so that $m_t^x = m_t^y = 0$ for all t. Substituting in (2.27) yields

$$m_t^z = m_0^z(1-2\mu)^t.$$

This coincides with the relaxation formula in the original Kac model. It is interesting to observe that the relaxation to equilibrium becomes slower for all the extensions that we have considered here.

3.2. Molecular chaos

The Stoßzahlansatz, or *repeated randomization* hypothesis consists in replacing the real dynamics by an effective dynamics in which the scattering mechanism is not kept fixed but gets replaced with an average. In this way, memory is being erased of where the scatterers are. An effective dynamics then works on the single-particle level and since the fraction of active scatterers equals μ, the one-site density matrix ν_t is either copied with probability $(1-\mu)$ or replaced with a scattered density matrix with probability μ. More precisely,

$$\Gamma^\star(\nu) \equiv (1-\mu)\nu + \mu U \nu U^\star \tag{3.1}$$

defines a completely positive map (the quantum equivalent of a discrete time stochastic dynamics, see [1]). The unitary U should be substituted from (2.1). On the level of the full density matrix ρ^N, we just make the product of copies of $\Gamma^\star$ over all N sites so that a product state is mapped onto a product state: if $\nu^N(j) \equiv \mathrm{Tr}_j[\rho^N]$, then

$$\Lambda^\star(\rho^N) \equiv \bigotimes_{j=1}^{N} \Gamma^\star(\nu^N(j)). \tag{3.2}$$

This is the dual of a map Λ and we write Λ_t for the map Λ applied t times. The effective dynamics thus amounts to replacing the time evolution $U_N(t)\cdot U_N(t)^\star$ (depending on the scatterers ϵ_j) by $\Lambda_t^\star \equiv \bigotimes \Gamma_t^\star$ (only depending on μ) and expectations at time t become

$$\mathrm{Tr}\left[\Lambda_t(A_N)\rho_0^N\right].$$

Computations are even simpler here.

The results of section 2.2 say essentially that the above effective dissipative dynamics reproduces the correct result (see lemma 4.2). For example,

$$\epsilon\text{-}\lim_{N\uparrow+\infty} \frac{1}{N}\,\mathrm{Tr}\left[M_N^z \rho_t^N\right] = \lim_{N\uparrow+\infty} \frac{1}{N}\,\mathrm{Tr}\left[\Lambda_t\left(M_N^z\right)\rho_0^N\right] = \mathrm{Tr}[\Gamma_t(\sigma_z)\nu_0]. \tag{3.3}$$

In other words, the unitary dynamics (2.2) (depending on the scatterers ϵ_j) is typically equivalent to the effective dynamics Γ_t on the one-particle level. The marginal ν_t of (2.22) exists and can be obtained from $\nu_{t+1} = \Gamma^\star \nu_t$. This specifies equation (2.27) and determines (2.30).

Note finally that the reduction of the unitary time evolution (2.2) to an effective dynamics on a one-particle system remains non-Abelian. Of course, our macroscopic observables start to commute,

$$\left[\frac{M_N^x}{N}, \frac{M_N^y}{N}\right] = \mathcal{O}\left(\frac{1}{N}\right)$$

but $\Gamma^\star$ remains defined on density matrices and does not yield a classical dynamics on diagonal elements of the density matrix ν_0.

As stated in theorem 2.1, as in the original Kac ring model (or as for the rigorous derivation of the Boltzmann equation), one really does not need this assumption of molecular chaos to get relaxation. It is replaced with statistical assumptions on the level of the initial conditions for the scatterers—they are typical with respect to the probability (2.15)—and for the spins—they are, e.g., randomly sampled with a fixed magnetization in (2.11). In this case we think of both the η (wavefunction) and the ϵ (classical scattering centres) as dynamical variables.

3.3. Autonomous equations and H-theorem

Equation (2.27) is autonomous in the sense that the value of $\vec{m}_t$ determines $\vec{m}_{t+1}$ once we know $\vec{h} \equiv h\vec{n} = (e_+ - e_-)\vec{n}/2$ in the Hamiltonian (2.4) and the fraction μ of scatterers:

$$\vec{m}_{t+1} = \vec{m}_t - 2\mu[(\vec{n}\times\vec{m}_t)\sin h\cos h - \vec{n}\times(\vec{n}\times\vec{m}_t)\sin^2 h]. \tag{3.4}$$

A more suggestive expression is obtained by decomposing $\vec{m}$ into a parallel and a perpendicular component along $\vec{n}$:

$$\vec{m} = (\vec{m}\cdot\vec{n})\vec{n} + \vec{n}\times(\vec{m}\times\vec{n}) \equiv \vec{m}_\parallel + \vec{m}_\perp$$

for which

$$\begin{aligned} \vec{m}_\parallel(t+1) &= \vec{m}_\parallel(t) \\ \vec{m}_\perp(t+1) &= (1-2\mu\sin^2 h)\vec{m}_\perp(t) + 2\mu\sin h\cos h\,\vec{m}_\perp(t)\times\vec{n}. \end{aligned} \tag{3.5}$$

The relaxation formula (2.27) simplifies for certain initial conditions. For example, when taking (2.11) or (2.12), so that $m_0^x = m_0^y = 0$, we get the damped oscillator

$$m_t^z = m_e^z + \left(m_0^z - m_e^z\right) r^t \cos\omega t$$

with $r^2 \equiv (1-\mu)^2 + \mu^2 + 2\mu(1-\mu)\cos 2h$ and $\tan(\omega) \equiv \mu\sin 2h/(1-\mu+\mu\cos 2h)$. The frequency is maximal for $h = \pi/4$ and the damping is maximal for $h = \pi/2$. In this case, the relaxation of m_t^z also looks autonomous but that is because of the special initial data.

In general, m_t^z does not relax autonomously and, in contrast with (2.29), its associated entropy

$$-\frac{1+m_t^z}{2}\log\frac{1+m_t^z}{2} - \frac{1-m_t^z}{2}\log\frac{1-m_t^z}{2}$$

need not be monotone (even in the case of (2.11) or (2.12)). Going to the bigger picture with three magnetization components reveals a new structure in which the entropy does increase. From (3.5), we see that $\|\vec{m}(t)\|$ decreases as

$$\|\vec{m}_\perp(t+1)\|^2 = (1-4\mu(1-\mu)\sin^2 h)\|\vec{m}_\perp(t)\|^2 \tag{3.6}$$

and hence, combined with (2.29), the monotonicity $s(t+1) \geqslant s(t)$ follows and strictly so if $\mu \in (0,1)$ and $h \neq k\pi$. We thus see here that also in the quantum case, autonomy on some macroscopic scale is intrinsically connected to monotonicity of the corresponding entropy.

3.4. Relaxation to equilibrium?

The relaxation can be read off clearly from (3.5): we get a spiral motion in the plane perpendicular to $\vec{n}$. The model does however not show the full glory of relaxation to equilibrium. This is already true in the original Kac ring model. As an example, consider a macroscopic variable which involves a two-spin function, in contrast to the single-site observables that we introduced in (2.5):

$$A_N = \sum_j \sigma_{z,j}\sigma_{z,j+1}.$$

It is easy to verify that A_N/N does not show relaxation. This is not surprising given the fact that two neighbouring spins live exactly the same history and thus they do not decorrelate. This seriously restricts the usefulness of the model for studies of relaxation but it remains possible to study the phenomenon of relaxation for the special macroscopic observables of (2.5).

On the other hand, the fact that we consider relaxation in terms of macroscopic observables should not be considered as a restriction. After all, equilibrium is characterized by a maximal entropy condition given macroscopic constraints. In particular, relaxation is not to be read off from the Liouville–von Neumann evolution for all microscopic details of the density matrix. The fact that we first let $N \uparrow +\infty$ (before time) avoids the presence of (quasi-)periodicities or Poincaré recurrence.

A final useful comparison concerns the notion of ergodicity which is quite universal, see, e.g., [5, 4]. Consider a quantum mechanical system with a discrete non-degenerate energy spectrum (E_n) with (ϕ_n) a complete set of orthonormal energy eigenfunctions. The wavefunction at time t is denoted by $\psi(x,t) = \exp[-\mathrm{i}Ht]\psi(x)$ for initial $\psi(x) = \sum_n c_n\phi_n(x)$. The time average of the expectation value of a Hermitian A is

$$\bar{A} \equiv \lim_{T\uparrow+\infty} \frac{1}{T}\int_0^T \mathrm{d}t \int \mathrm{d}x\ \psi^\star(x,t)\, A\psi(x,t).$$

It is rather easy to see in this case (but it remains true in a much broader context) that

$$\bar{A} = \sum_n |c_n|^2 \int \mathrm{d}x\ \phi_n^\star(x)\, A\phi_n(x) \tag{3.7}$$

which could be argued to correspond to the microcanonical ensemble. Note however that no use was made in the above of the fact that N is large nor of the fact that we consider macroscopic variables. The result (3.7) is equally not very satisfactory: we are interested in the typical manifest behaviour over realistic time spans for large systems while (3.7) gives trivial information about an infinite time average that cannot be identified with relaxation phenomena. There is in fact no relaxation to equilibrium on a microscopic scale even though (3.7) always holds.

3.5. Higher spins

Instead of considering spin-$\frac{1}{2}$ particles, it is also possible to set up exactly the same problem as above for a Hilbert space $\mathcal{H}_N = \bigotimes_{j=1}^N \mathbb{C}_j^n$. We then deal with $n \times n$ matrices but there is little difference in the computations except for one extra complication: there is a larger class of conserved quantities. To explain this, we split the single-site observable A (a Hermitian matrix in $\mathbb{C}^{n\times n}$) into a part that commutes with the Hamiltonian H, and its orthogonal complement:

$$A = A_H \oplus A_\perp \tag{3.8}$$

with $[A_H, H] = 0$ and with respect to the scalar product

$$\langle A|B\rangle = \mathrm{Tr}[A^\star B]. \tag{3.9}$$

The set of commuting observables $\mathcal{A}_H$ is a vector space with dimension

$$\dim(\mathcal{A}_H) = n.$$

If $n > 2$ there is more than just the $\mathbb{1}$ and $H \in \mathcal{A}_H$ (e.g., not all functions of H can be expressed as linear combinations of $\mathbb{1}$ and H).

Because of the triviality of the dynamics as was discussed in the previous section, these extra conserved quantities give rise to conserved quantities for the full system dynamics and none of them disappears as the size of the system increases. For more realistic systems, truly interacting, we expect that, in the thermodynamic limit, essentially only such physical quantities as particle number and total energy remain conserved. In that respect, the case $n = 2$ is more physical and that is the reason why we have concentrated on it from the beginning. The mathematical theory can however be completed if we change the definition of the microcanonical ensemble (2.7) to include the extra conserved quantities.

4. Proof of results

Since it is important for physical interpretation that the relaxation curve (2.27) is obtained for almost all choices of scatterers ϵ drawn from (2.14), we must first recall in lemma 4.1 the

thermodynamic equivalence with a grand-canonical set-up. Next, in lemma 4.2 it is shown that an effective dynamics reproduces the time evolution of our macroscopic observables. Using this, and from lemma 4.3, the propagation of condition (2.21) is obtained. It is therefore sufficient to derive (2.27) to get (2.26). This computation is introduced by lemma 4.4. The entropy (2.29) and its monotonicity can then also be computed using the obtained law of large numbers.

Let $\Omega \equiv \{0, 1\}$ and $\epsilon \in \Omega^N$. We have already defined the 'canonical' measure $\mathbb{P}_\mu^N$ in (2.15) on Ω_μ^N of (2.14). Define now its 'grand-canonical' version, the probability measure $\mathbb{P}_\mu$ on $\Omega^{\mathbb{N}}$, as the Bernoulli measure with density $\mathbb{P}_\mu(\epsilon_j = 1) = \mu$. $\mathbb{E}_\mu^N(\cdot)$ and $\mathbb{E}_\mu(\cdot)$ denote the expectation values with respect to $\mathbb{P}_\mu^N$ and $\mathbb{P}_\mu$, respectively. The following equivalence is standard.

Lemma 4.1. *For all functions* $g \equiv g(\epsilon_1, \ldots, \epsilon_k)$ *on* Ω^k

$$\left|\mathbb{E}_\mu^N(g) - \mathbb{E}_\mu(g)\right| \leqslant \frac{2^{2k+1}k}{N}\|g\| \qquad if \quad 2k \leqslant N$$

with

$$\|g\| \equiv \max_{\xi \in \Omega^k} |g(\xi)|.$$

Proof.

$$\left|\mathbb{E}_\mu^N(g) - \mathbb{E}_\mu(g)\right| \leqslant \sum_{\xi \in \Omega^k} |g(\xi)| \left|\mathbb{P}_\mu^N(\xi) - \mathbb{P}_\mu(\xi)\right|.$$

The right-hand side can be bounded by developing $\mathbb{P}_\mu^N(\xi)$ in conditional expectations. By explicit calculation

$$\left|\mathbb{P}_\mu^N(\xi_{k+1} = 1 \mid (\xi_1, \ldots, \xi_k) = \xi) - \mu\right| \leqslant \frac{k+1}{N-k}$$

after which a simple calculation finishes the proof. □

We define for fixed j and matrix element per element

$$\Gamma_t(A_j) \equiv \mathbb{E}_\mu A_j(t, \epsilon) = \underset{N\uparrow+\infty}{\epsilon\text{-}\lim} \mathbb{E}_\mu^N A_j(t, \epsilon)$$

where the last equality follows from lemma 4.1.

Lemma 4.2. *For macroscopic observables* $A_N = \sum_j A_j$ *and continuous functions* f*, the dynamics is typically equivalent to the semigroup dynamics* Λ_t*, i.e.,*

$$\underset{N\uparrow+\infty}{\epsilon\text{-}\lim} \operatorname{Tr}\left[f\left(\frac{A_N}{N}\right)\rho_t^N\right] = \lim_{N\uparrow+\infty} \operatorname{Tr}\left[f\left(\Lambda_t\left(\frac{A_N}{N}\right)\right)\rho_0^N\right].$$

Proof. The proof is for all monomials. By the Stone–Weierstrass theorem, this proves the theorem for all continuous functions f.

Let

$$D(n) \equiv \{1, \ldots, N\}^n \qquad T(n, t) \equiv \{K \in D(n) | \forall i, j \leqslant n : i \neq j \Rightarrow |K_i - K_j| > t\}$$

and

$$A(K) \equiv A_{j(1)}^{k(1)} \otimes \cdots \otimes A_{j(l)}^{k(l)}$$

where

$$K = \{\underbrace{j(1), \ldots, j(1)}_{k(1)}, \ldots, \underbrace{j(l), \ldots, j(l)}_{k(l)}\}.$$

We also need the index set $J(K) \equiv \{j | \exists p : K_p = j\}$.

We begin by calculating the ϵ-average. Denote $A_N(t,\epsilon) \equiv U_N(t)^\star A_N U_N(t)$, and similarly for the one-site observables A_j:

$$\operatorname{Tr}\left[\left(\frac{A_N}{N}(t,\epsilon)\right)^n \rho_0^N\right] = \frac{1}{N^n}\sum_w \sum_{K\in T(n,t)} \prod_{j\in J(K)}^{n} \operatorname{Tr}\left[A_j(t,\epsilon)\chi_{j,w}^N\right] \prod_{j\notin J(K)} \operatorname{Tr}\left[\chi_{j,w}^N\right] + \frac{1}{N^n}\sum_{K\in D(n)\setminus T(n,t)} \operatorname{Tr}\left[A(K)(t,\epsilon)\rho_0^N\right].$$

The last term can be dropped because

$$\left|\frac{1}{N^n}\sum_{K\in D(n)\setminus T(n,t)} \operatorname{Tr}\left[A(K)(t,\epsilon)\rho_0^N\right]\right| \leqslant \frac{1}{N^n}\sum_{K\in D(n)\setminus T(n,t)} \|A(K)\| = \mathcal{O}\left(\frac{\|A\|^n}{N}\right).$$

We apply lemma 4.1 with $g_{K,w}(\epsilon) \equiv \prod_{j\in K} \operatorname{Tr}\left[A_j(t,\epsilon)\chi_{j,w}^N\right]$ as a function on Ω^{tn}. For $K \in T(n,t)$,

$$\left|\mathbb{E}_\mu^N \prod_{j\in K} \operatorname{Tr}\left[A_j(t,\epsilon)\chi_{j,w}^N\right] - \left(\mathbb{E}_\mu^C \operatorname{Tr}\left[A_j(t,\epsilon)\chi_{j,w}^N\right]\right)^n\right| \leqslant \frac{2^{2tn}2(tn)}{N} C^n \|A\|^n.$$

We replace $\mathbb{E}_\mu \operatorname{Tr}\left[A_j(t,\epsilon)\chi_{j,w}^N\right]$ by $\operatorname{Tr}\left[\Gamma_t(A_j)\chi_{j,w}^N\right]$ and add terms of the order of N^{-1} (essentially the ones we first subtracted) to arrive at the average from the lemma.

For the variance, we similarly write

$$\mathbb{E}_\mu^N\left(\operatorname{Tr}\left[\left(\frac{A_N}{N}(t,\epsilon)\right)^n \rho_0^N\right]\right)^2$$

as

$$\frac{1}{N^{2n}}\sum_{K,L\in T(n,t)} \mathbb{E}_\mu^N\left(\operatorname{Tr}\left[A(K)(t,\epsilon)\rho_0^N\right] \operatorname{Tr}\left[A(L)(t,\epsilon)\rho_0^N\right]\right) + \mathcal{O}(N^{-1}).$$

Also the contribution of sets K and L that are within a distance less than t is of lower order. We apply lemma 4.1 just as above. □

Lemma 4.3. *For all continuous functions f and macroscopic observables A_N,*

$$\epsilon\text{-}\lim_{N\uparrow+\infty} \operatorname{Tr}\left[f\left(\frac{A_N}{N}\right)\rho_t^N\right] = \lim_{N\uparrow+\infty} f\left(\operatorname{Tr}\left[\Lambda_t\left(\frac{A_N}{N}\right)\rho_0^N\right]\right)$$

at all times t.

Proof. By lemma 4.2 we can insert the initial density matrix. The statement then follows from applying condition (2.21) to the observable $\sum_j \Gamma_t(A_j)$. □

As an illustration of the above lemma, we show that it can be applied for the state ρ_0^N given by (2.11).

Example. Call P_a^N the projector on the eigenvalue a of the macroscopic variable A_N/N. Let I^N be the spectrum of A_N/N and $I = \overline{\cup I^N}$. We want to establish a large deviation property for the probability measures W^N:

$$W^N(K) \equiv \sum_{a \in I^N \cap K} \mathrm{Tr}\left[P_a^N \rho_0^N\right]. \tag{4.1}$$

The product $P_a^N \rho_0^N$ can be written as a sum over one-dimensional projections and (4.1) can be developed and written in terms of

$$\mathrm{Tr}\left[P_{a_j} Q_{m_j}\right]$$

which takes four possible values depending on (p_j, m_j). The problem amounts therefore to estimating the probability that at the same time $\sum_j p_j$, $\sum_j m_j$ and $\sum_j m_j p_j$ take on specific values when the (p_j, m_j) are independent and identically distributed with uniform weights. This can be found from the multinomial distribution and Stirling's formula gives the large deviation rate function corresponding to (4.1). As a consequence, a law of large numbers is satisfied and in the thermodynamic limit the state (2.11) concentrates on only one eigenvalue of A_N/N.

We now come back to (2.22) defining the marginal ν_0. We split the Pauli matrices $\sigma_\alpha = \sigma_H^\alpha + \sigma_\perp^\alpha$ as in (3.8) and parametrize in the single-site Hamiltonian basis:

$$\nu_0 = \begin{pmatrix} p & \beta \\ \bar{\beta} & 1-p \end{pmatrix} \qquad \sigma_\perp^\alpha = \begin{pmatrix} 0 & \theta_\alpha \\ \bar{\theta}_\alpha & 0 \end{pmatrix}. \tag{4.2}$$

The typical value of the magnetization is defined to be

$$\vec{m}_t \equiv \epsilon\text{-}\lim_{N\uparrow+\infty} \mathrm{Tr}\left[\frac{\vec{M}_N}{N}\rho_t^N\right].$$

Lemma 4.4.

$$\vec{m}_t = \vec{m}_e + \mathrm{Re}\left[\left(\vec{m}_0 - \vec{m}_e + \frac{\mathrm{i}(\vec{h}\times\vec{m}_0)}{h}\right)(1-\mu+\mu\exp[\mathrm{i}(e_+ - e_-)])^t\right]. \tag{4.3}$$

Proof. Fixing t, we write $k(\epsilon) \equiv \sum_{j=1}^t \epsilon_j$ and obtain

$$\Gamma_t(\sigma_\perp^\alpha) = \mathbb{E}_\mu[U^{k(\epsilon)^\star}\sigma_\perp^\alpha U^{k(\epsilon)}] = \begin{pmatrix} 0 & z_\alpha(t) \\ \bar{z}_\alpha(t) & 0 \end{pmatrix}$$

with

$$z_\alpha(t) \equiv \theta_\alpha\,(1-\mu+\mu\exp[\mathrm{i}(e_+ - e_-)]). \tag{4.4}$$

σ_H^α is invariant under the time evolution and it is easy to verify that

$$\lim_{t\uparrow+\infty} \mathrm{Tr}[\Gamma_t(\sigma_\alpha)\nu_0] = \mathrm{Tr}\left[\sigma_H^\alpha \nu_0\right] = m_e^\alpha.$$

This yields

$$m_t^\alpha = m_e^\alpha + \mathrm{Re}[2\bar{\beta}\theta_\alpha(1-\mu+\mu\exp[\mathrm{i}(e_+ - e_-)])^t].$$

Now it is a matter of writing $2\bar{\beta}\theta_\alpha$ as

$$2\bar{\beta}\theta_\alpha = \mathrm{Tr}[\nu_0\sigma_\perp^\alpha] - \frac{\mathrm{Tr}[H\nu_0\sigma_\perp^\alpha]}{h}$$

then inserting $H = \vec{h} \cdot \vec{\sigma}$ and recognizing the resulting complex numbers as components of the vector product $\vec{h} \times \vec{m}_0$, to find (4.3) almost immediately. □

The asymptotics (2.28) is trivial and the proof of theorem 2.1 is thus completed if we add equation (2.29) for the entropy $\mathcal{S}[\nu_t]$.

Proof. The explicit formula (2.29) can immediately be obtained by using the Bloch representation:

$$\nu_t = \frac{\mathbb{1} + \vec{m}_t \cdot \vec{\sigma}}{2}.$$

Inserting it in the entropy density (2.30) yields the result.

Concerning the increase of entropy, denote by ν_e the one-site marginal of the equilibrium state (2.6), with magnetization $\vec{m}_e$. A short computation gives

$$\nu_e = \begin{pmatrix} p & 0 \\ 0 & 1-p \end{pmatrix}$$

where p is the first diagonal element of ν_0, see (4.2).

The relative entropy between two marginals is defined as

$$\mathcal{S}[\nu|\nu_e] = \mathrm{Tr}[\nu \log \nu] - \mathrm{Tr}[\nu \log \nu_e]. \tag{4.5}$$

We immediately have

$$\mathcal{S}[\nu_t|\nu_e] = \mathcal{S}[\nu_e] - \mathcal{S}[\nu_t]$$

since the Γ-dynamics works only on the off-diagonal elements of ν_0. Together with the entropy-contraction inequality for completely positive maps

$$\mathcal{S}[\Gamma^\star(\nu_t)|\Gamma^\star(\nu_e)] \leqslant \mathcal{S}[\nu_t|\nu_e]$$

and since $\Gamma^\star(\nu_e) = \nu_e$, we find

$$\mathcal{S}[\Gamma^\star(\nu_t)] = \mathcal{S}[\nu_{t+1}] \geqslant \mathcal{S}[\nu_t].$$

□

Acknowledgments

We are grateful to I Callens, A van Enter, B Haegeman and S Pirogov for useful discussions.

References

[1] Alicki R and Fannes M 2001 *Quantum Dynamical Systems* (Oxford: Oxford University Press)
[2] Bricmont J 1995 Science of chaos or chaos in science? (appendix 1) *Physicalia* **17** 159–208
[3] Callens I, De Roeck W, Jacobs T, Netočný K and Maes C 2002 Quantum entropy production as a measure for irreversibility *Preprint* cond-mat/0211252 (*Physica D* at press)
[4] Evans D E 1977 Irreducible quantum dynamical semigroups *Commun. Math. Phys.* **54** 293–7
[5] Frigerio A 1978 Stationary states of quantum dynamical semigroups *Commun. Math. Phys.* **63** 269–76
[6] Kac M 1959 *Probability and Related Topics in Physical Sciences* (New York: Interscience) chapter 14
[7] Robinson D W 1973 Return to equilibrium *Commun. Math. Phys.* **31** 171–89
[8] Sewell G L 2002 *Quantum Mechanics and Its Emergent Macrophysics* (Princeton, NJ: Princeton University Press)
[9] Thompson C J 1972 *Mathematical Statistical Mechanics* (Princeton, NJ: Princeton University Press)