USING DENSITY MATRICES
IN CLASSICAL DYNAMICAL SYSTEMS

MARK FANNES

Instituut Theoretische Fysica, K.U. Leuven
Celestijnenlaan 200D, B-3001 Heverlee, Belgium
E-mail: mark.fannes@fys.kuleuven.ac.be

1. Introduction. The aim of this paper is to show how typically quantummechanical objects, such as operators and density matrices, can be used for studying classical dynamical systems. We will in particular consider statistical descriptions of dynamical systems in discrete time, where one tries to study the actual system in terms of a restricted class of model systems, in our case the non-commutative shifts on half-chains of quantum spins. This is quite different from the Koopman formalism, where one constructs a quantummechanical evolution on the Hilbert space $\mathfrak{H}$ of square integrable functions on the phase space in such a way that the evolution of the multiplication operators on $\mathfrak{H}$ coincides with the classical evolution of the phase space functions.

There are several reasons to consider such descriptions. Although one may very well stay with classical statistical models, i.e. shifts on symbolic sequences, models in terms of non-commutative shifts may be quite efficient and tractable. A second reason to consider such models is to put classical dynamical systems and their statistical description in the wider non-commutative setting. Finally, this type of models leads to an interesting class of shift-invariant states on quantum spin chains which might be worthwhile to study for its own sake.

2. Classical statistical models. [3,5] We will describe a classical dynamical system by a continuous transformation θ of a compact metrizable set X: the phase space of the system. The map θ corresponds to the evolution of the system during a single time step. The dynamics is called reversible if θ has a continuous inverse. More precisely, the dynamics is given by the discrete group $\{\theta^n \mid n \in \mathbf{Z}\}$ or semigroup $\{\theta^n \mid n \in \mathbf{N}\}$ of continuous transformations of X according to whether θ is invertible or not. The last

ingredient in the specification of a dynamical system is the invariant measure μ which is a regular probability measure on X such that $\mu(A) = \mu(\theta^{-1}(A))$ for any Borel subset A of X.

The basic building blocks for constructing classical statistical models of (X, θ, μ) are the finite probability spaces $(\mathbf{Z}_q, \lambda)$: $\mathbf{Z}_q = \{0, 1, \ldots, q-1\}$ and $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_{q-1})$ with $\lambda_j \geq 0$ and $\lambda_0 + \lambda_1 + \cdots + \lambda_{q-1} = 1$. We call $(\mathbf{Z}_q, \lambda)$ a "coarse-grained" picture of (X, μ) if there is a mapping γ from $\mathbf{Z}_q$ to the Borel subsets $\mathfrak{B}(X)$ of X such that

    i) $\gamma(j) \cap \gamma(k) = \emptyset$ if $j \neq k$,
    ii) $\bigcup_j \gamma(j) = X$ and
    iii) $\lambda_j = \mu(\gamma(j))$.

$\{\gamma(j) \mid j = 0, 1, \ldots, q-1\}$ is called a measurable size-q partition of X. Successive applications of the dynamics on X, up to time T, will refine partitions and produce ever more precise coarse-grained pictures $(\prod_{t=0}^{T-1} \mathbf{Z}_q, \lambda^{(T)})$ of the dynamical system given in terms of maps $\gamma^{(T)}$. Let $\mathbf{j} = (j_0, j_1, \ldots, j_{T-1})$ with $j_t \in \mathbf{Z}_q$ be a configuration of length T, then $\gamma^{(T)}$ is given by:

$$\gamma^{(T)} : \underbrace{\mathbf{Z}_q \times \mathbf{Z}_q \times \cdots \times \mathbf{Z}_q}_{T-\text{times}} \to \mathfrak{B}(X) : \mathbf{j} \mapsto \bigcap_{t=0}^{T-1} \theta^{-t}(\gamma(j_t)).$$

The mappings $\gamma^{(T)}$ are compatible in the sense that

$$\bigcup_{j'} \gamma^{(T+1)}(\mathbf{j} \times (j')) = \gamma^{(T)}(\mathbf{j}).$$

In this formula $\mathbf{j} \times \mathbf{j}'$ denotes the concatenation of the configurations $\mathbf{j}$ and $\mathbf{j}'$. We can therefore glue the different $\gamma^{(T)}$ together to obtain a global map $\overline{\gamma} : (\mathbf{Z}_q)^{\mathbf{N}} \to \mathfrak{B}(X)$ which is the limit of the $\gamma^{(T)}$. Furthermore

$$\bigcup_{j'} \overline{\gamma}((j') \times \mathbf{j}) = \theta^{-1}(\overline{\gamma}(\mathbf{j})).$$

This last relation, together with the time-invariance of μ, implies that the measure $\overline{\lambda} := \mu \circ \overline{\gamma}$ is invariant for the left-shift σ on the half-infinite chain $(\mathbf{Z}_q)^{\mathbf{N}}$ and that $\theta \circ \overline{\gamma} = \overline{\gamma} \circ \sigma$.

3. Quantum statistical models. [1,2] We will now modify the scheme of above essentially by replacing the classical finite probability space $(\mathbf{Z}_q, \lambda)$ and the coarse-graining map γ by suitable non-commutative versions. In order to do so, we must first lift the formalism to an algebraic level, i.e. describe the classical dynamical system in terms of

    i) the algebra $\mathcal{C}(X)$ of continuous complex-valued functions on phase space,
    ii) the single-step automorphism or homomorphism $\Theta(f) := f \circ \theta$, and
    iii) the functional $\mu(f) := \int_X d\mu \, f$.

We then replace the couple $(\mathbf{Z}_q, \lambda)$ by $(\mathcal{M}_q, \rho)$ where $\mathcal{M}_q$ are the $q \times q$ complex matrices and where ρ is a density matrix on $\mathbf{C}^q$. It remains to specify the coarse-graining map Γ and how it refines under the dynamics. Following an idea of Lindblad's [4], we put

$$\Gamma([a_{k\ell}]_{k\ell}) := \sum_{k\ell} a_{k\ell} \, \overline{f_k} \, f_\ell,$$

where $\mathcal{F} = \{f_0, f_1, \ldots, f_{q-1}\}$ is an "operational partition of unity", i.e. a collection of functions in $\mathcal{C}(X)$ such that $\sum_k |f_k|^2 = 1$. In the sequel, such $\mathcal{F}$ will simply be called partitions. The expression for a refined mapping $\Gamma^{(T)}$ is similar to that of Γ except that k has to be replaced by a T-tuple $\mathbf{k} = \{k_0, k_1, \ldots, k_{T-1}\}$ and f_k by

$$f_{\mathbf{k}} := f_{k_0} \, \Theta(f_{k_1}) \cdots \Theta^{T-1}(f_{k_{T-1}}).$$

The $\Gamma^{(T)}$ are completely positive, normalized and right-compatible in the sense that

$$\Gamma^{(T+1)}(A \otimes \mathbb{1}) = \Gamma^{(T)}(A), \qquad A \in \mathcal{M}_q^{\otimes T}.$$

They define therefore a global statistical picture $\overline{\Gamma}$ of the dynamical system in terms of a half-chain of quantum spins

$$\overline{\Gamma} : \mathcal{M}_q^{\otimes \mathbf{N}} \to \mathcal{C}(X) : \overline{\Gamma} := \lim \Gamma^{(T)}.$$

Denoting by Σ the right-shift $A \mapsto \mathbb{1} \otimes A$ on $\mathcal{M}_q^{\otimes \mathbf{N}}$ we find that $\overline{\Gamma} \circ \Sigma = \Theta \circ \overline{\Gamma}$. Again, by time invariance of μ, the global state $\overline{\rho} := \mu \circ \overline{\Gamma}$ is shift-invariant on the quantum half-chain.

It is also possible, and useful, to consider the larger algebra $\mathcal{L}^\infty(X, \mu)$ instead of $\mathcal{C}(X)$ and to extend the dynamics Θ and the functional μ to this algebra in the obvious way. The notion of operational partition of unity in continuous functions clearly extends to the measurable functions and so does the construction of the statistical models. We can moreover associate in a natural way to a partition $\mathcal{C} = \{C_0, C_1, \ldots, C_{q-1}\}$ of the phase space X into disjoint measurable subsets the measurable partition of unity $\{\chi_{C_0}, \chi_{C_1}, \ldots, \chi_{C_{q-1}}\}$ where χ_A is the characteristic function of the set A. By this association, intersection of subsets of X corresponds to multiplication of their characteristic functions, which means that the notions of refinement in the classical sense and in the sense of quantum statistical models agree. It is quite obvious that a partition of X into disjoint subsets leads to consider diagonal density matrices on finite portions of the spin half-chain. More general operational partitions will usually produce highly non-diagonal density matrices that don't obey any particular commutation relations. The state $\overline{\rho}$ has however a particular structure: it is completely non-entangled:

PROPOSITION 1. *Let* $\{e_0, e_1, \ldots e_{q-1}\}$ *denote the canonical basis of* $\mathbf{C}^q$ *and* $\mathrm{P}_{\mathcal{F}}$ *the projection-valued function*

$$\mathrm{P}_{\mathcal{F}} := |\sum_k f_k \, e_k\rangle\langle\sum_k f_k \, e_k|$$

on X, *then*

$$\overline{\rho} = \int_X d\mu \, \otimes_{j \in \mathbf{N}} \Theta^j(\mathrm{P}_{\mathcal{F}}).$$

We can of course rewrite the state $\overline{\rho}$ in terms of a stationary, discrete time, stochastic dynamics on the closed subset $\mathrm{P}_0 := \mathrm{P}_{\mathcal{F}}(X)$ of the pure state space P of $\mathcal{M}_q$. We attribute the probability

$$\nu(\{\pi(t) \in A_t\}) := \mu\left(\bigcap_t \theta^{-t}\left(\mathrm{P}_{\mathcal{F}}^{-1}(A_t)\right)\right)$$

to the paths π in P_0 that cross A_t at time t, where A_t is a measurable subset of P_0:

$$\overline{\rho} = \int d\nu(\pi) \, \otimes_{t \in \mathbf{N}} \pi(t).$$

Except for the purity condition on the states on $\mathcal{M}_q$, we obtain in this way the full class of classical, shift-invariant states on $\mathcal{M}_q^{\otimes \mathbf{N}}$:

$$\int d\nu(\varsigma) \otimes_{t \in \mathbf{N}} \varsigma(t),$$

where ν is a stationary, discrete time, stochastic process on the state space $\mathcal{S}$ of $\mathcal{M}_q$.

Let us choose two strictly local observables A and B in $\mathcal{M}_q^{\otimes \mathbf{N}}$. If we shift B far enough to the right, we can ensure that A and $\Sigma^\ell(B)$ are localized on disjoint subsets of $\mathbf{N}$. In this case:

$$\overline{\rho}(A \, \Sigma^\ell(B)) = \int_X d\mu \, a \, \Theta^\ell(b),$$

where a is the function $x \in X \mapsto \mathrm{Tr} \otimes_t \Theta^t(\mathrm{P}_{\mathcal{F}}) \, A$ and b is defined in a similar way. Therefore, the models of an ergodic dynamical system will be ergodic with respect to the shift and the models of a mixing system will be mixing. Conversely, a dynamical system will be ergodic (or mixing) only if all of its models are ergodic (mixing).

4. Computing Renyi entropies. As an example, we will compute the second mean Renyi entropy r_2 for the model of an ergodic dynamical system determined by the partition $\mathcal{F} = \{f_0, f_1, \ldots, f_{q-1}\}$. The mean k^{th} order Renyi entropy $r_k(\omega)$ of a shift-invariant state ω on $\mathcal{M}_q^{\otimes \mathbf{N}}$ is defined as:

$$r_k(\omega) := \limsup_{T \to \infty} \frac{-1}{T(k-1)} \log \mathrm{Tr}\,(\omega_T^k).$$

In this formula, ω_T is the reduced density matrix of ω for the first T sites of the half-chain. In general, it is not possible to replace "limsup" by "lim".

Applying the construction of r_2 to our model state $\overline{\rho}$, we find:

$$
\begin{aligned}
r_2(\overline{\rho}) &= \limsup_{T \to \infty} \frac{-1}{T} \log \mathrm{Tr} \left(\int_X d\mu \bigotimes_{t=0}^{T-1} \Theta^t(\mathrm{P}_{\mathcal{F}}) \right)^2 \\
&= \limsup_{T \to \infty} \frac{-1}{T} \log \int_{X \times X} d\mu(x_1) d\mu(x_2) \prod_{t=0}^{T-1} \mathrm{Tr}\, \mathrm{P}_{\mathcal{F}}(\theta^t(x_1)) \, \mathrm{P}_{\mathcal{F}}(\theta^t(x_2)) \\
&= \limsup_{T \to \infty} \frac{-1}{T} \log \int_{X \times X} d\mu(x_1) d\mu(x_2) \prod_{t=0}^{T-1} \left| \sum_{j=0}^{q-1} \overline{f_j(\theta^t(x_1))} \, f_j(\theta^t(x_2)) \right|^2. \quad (*)
\end{aligned}
$$

At this point, we are unable to continue the computation unless we make some assumption about the product dynamics.

Let us first suppose that $(X \times X, \theta \times \theta, \mu \times \mu)$ is also ergodic. Writing

$$\prod_{t=0}^{T-1} \left| \sum_{j=0}^{q-1} \overline{f_j(\theta^t(x_1))} \, f_j(\theta^t(x_2)) \right|^2 = \exp\left(2 \sum_{t=0}^{T-1} \log \left| \sum_{j=0}^{q-1} \overline{f_j(\theta^t(x_1))} \, f_j(\theta^t(x_2)) \right| \right),$$

and using that for almost any x_1 and x_2

$$\lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \log \left| \sum_{j=0}^{q-1} \overline{f_j(\theta^t(x_1))} \, f_j(\theta^t(x_2)) \right| = \int_{X \times X} d\mu(y_1) d\mu(y_2) \log \left| \sum_{j=0}^{q-1} \overline{f_j(y_1)} \, f_j(y_2) \right|,$$

we find that r_2 exists as a limit and that it is explicitly given by

$$r_2(\overline{\rho}) = -2 \int_{X \times X} d\mu(y_1) d\mu(y_2) \log \left| \sum_{j=0}^{q-1} \overline{f_j(y_1)} f_j(y_2) \right|.$$

If the product dynamics is non-ergodic, the resulting entropy might be quite different. Consider e.g. the rotation of the unit circle S_1 in $\mathbf{R}^2$ over an angle α_0 which is an irrational multiple of π. The uniform measure $d\alpha/2\pi$ is ergodic. The product system however, has a conserved quantity:

$$(\alpha_1, \alpha_2) \mapsto (\alpha_1 - \alpha_2) \bmod 2\pi$$

and is therefore non-ergodic. Let us now resume the computation of r_2 at $(*)$:

$$r_2(\overline{\rho}) = \limsup_{T \to \infty} \frac{-1}{T} \log \frac{1}{4\pi^2} \int_0^{2\pi} d\alpha_1 \int_0^{2\pi} d\alpha_2 \prod_{t=0}^{T-1} \left| \sum_{j=0}^{q-1} \overline{f_j(\alpha_1 - t\alpha_0)} f_j(\alpha_2 - t\alpha_0) \right|^2$$

$$= \limsup_{T \to \infty} \frac{-1}{T} \log \frac{1}{4\pi^2} \int_0^{2\pi} d\beta \int_0^{2\pi} d\alpha \prod_{t=0}^{T-1} \left| \sum_{j=0}^{q-1} \overline{f_j(\alpha - t\alpha_0)} f_j(\alpha + \beta - t\alpha_0) \right|^2.$$

Using the ergodicity of an irrational shift on the circle,

$$\lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \log \left| \sum_{j=0}^{q-1} \overline{f_j(\alpha - t\alpha_0)} f_j(\alpha + \beta - t\alpha_0) \right| = \frac{1}{2\pi} \int_0^{2\pi} d\alpha \log \left| \sum_{j=0}^{q-1} \overline{f_j(\alpha)} f_j(\alpha + \beta) \right|.$$

By Laplace's formula we eventually obtain

$$r_2(\overline{\rho}) = \limsup_{T \to \infty} \frac{-1}{T} \log \frac{1}{2\pi} \int_0^{2\pi} d\beta \exp \left(2T \frac{1}{2\pi} \int_0^{2\pi} d\alpha \log \left| \sum_{j=0}^{q-1} \overline{f_j(\alpha)} f_j(\alpha + \beta) \right| \right)$$

$$= -2 \sup_{0 \le \beta < 2\pi} \frac{1}{2\pi} \int_0^{2\pi} d\alpha \log \left| \sum_{j=0}^{q-1} \overline{f_j(\alpha)} f_j(\alpha + \beta) \right| = 0,$$

because $\mathcal{F}$ is a partition of unity.

THEOREM 1. *Suppose that (X, θ, μ) is mixing, that $\mathcal{F} = \{f_0, f_1, \ldots, f_{q-1}\}$ is a partition of unity and that $\overline{\rho}$ denotes the shift-invariant state induced by $\mathcal{F}$ on $\mathcal{M}_q^{\otimes \mathbf{N}}$. For any $k \in \{2, 3, \ldots\}$, the kth mean Renyi entropy $r_k(\overline{\rho})$ exists as a limit and is given by*

$$r_k(\overline{\rho}) = \frac{-k}{2(k-1)} r_2(\overline{\rho}) = \frac{-k}{k-1} \int_{X \times X} d\mu(y_1) d\mu(y_2) \log \left| \left(\sum_{j=0}^{q-1} \overline{f_j(y_1)} f_j(y_2) \right) \right|.$$

P r o o f. Because the system is mixing, arbitrary products of copies of the dynamical system are mixing too. We can therefore repeat for general k, the computation of $r_2(\overline{\rho})$ of above. As $\overline{\rho}$ is a convex combination of pure product states, we can easily compute the k-fold traces in the expression for r_k and reach the statement of the theorem. ∎

The computation above is an example where the "replica trick" can be explicitly performed. It is rather obvious that, in this case, the computation of the higher Renyi entropies doesn't contain any valuable information. An "analytic continuation" of the formula for $k \to 1$ would in fact suggest an infinite von Neumann mean entropy for the

states that we consider. This is impossible however as the mean von Neumann entropy of a partition of size q cannot exceed $\log q$.

5. The von Neumann mean entropy. We will show in this section that the usual entropy density $\mathsf{h}(\mathcal{F})$ of a statistical model $\overline{\rho}$, induced by a partition $\mathcal{F}$, is bounded by the Kolmogorov-Sinai invariant. It is unclear for the moment whether a closed expression for $\mathsf{h}(\mathcal{F})$ can be obtained.

THEOREM 2. *Let $\mathcal{F} = \{f_0, f_1, \ldots, f_{q-1}\}$ be a measurable partition of unity of the dynamical system (X, θ, μ) and $\overline{\rho}$ the corresponding state on $\mathcal{M}_q^{\otimes \mathbf{N}}$, then the mean von Neumann entropy $\mathsf{h}(\mathcal{F})$ of $\overline{\rho}$ is bounded from above by the Kolmogorov-Sinai invariant h^{KS} of (X, θ, μ).*

P r o o f. The idea of the proof is to partition, up to time T, the phase space X into sufficiently fine subsets so that $\otimes_{t=0}^{T-1} \Theta^t(\mathrm{P}_\mathcal{F})$ is almost constant on such a subset.

Choose $\epsilon > 0$ and let $\mathcal{C} = \{C_0, C_1, \ldots, C_{p-1}\}$ be a partition of X into disjoint measurable subsets such that the variation of $\mathrm{P}_\mathcal{F}$ over each C_j doesn't exceed ϵ:

$$\sup_{x,y \in C_j} \|\mathrm{P}_\mathcal{F}(x) - \mathrm{P}_\mathcal{F}(y)\| \leq \epsilon \qquad j = 0, 1, \ldots, p-1.$$

The usual entropy function $x \in [0,1] \mapsto -x \log x$ will be denoted by η. If we put for a T-tuple $\mathbf{j} = (j_0, j_1, \ldots, j_T)$ $C_{\mathbf{j}} = \cap_t \theta^{-t}(C_{j_t})$, then

$$S(\overline{\rho}_T) = S\Big(\int_X d\mu \otimes_{t=0}^{T-1} \Theta^t(\mathrm{P}_\mathcal{F})\Big) = S\Big(\sum_{\mathbf{j}} \int_{C_{\mathbf{j}}} d\mu \otimes_{t=0}^{T-1} \Theta^t(\mathrm{P}_\mathcal{F})\Big)$$

$$\leq \sum_{\mathbf{j}} \eta(\mu(C_{\mathbf{j}})) + \sum_{\mathbf{j}} \mu(C_{\mathbf{j}}) \, S\Big(\frac{1}{\mu(C_{\mathbf{j}})} \int_{C_{\mathbf{j}}} d\mu \otimes_{t=0}^{T-1} \Theta^t(\mathrm{P}_\mathcal{F})\Big)$$

$$\leq \sum_{\mathbf{j}} \eta(\mu(C_{\mathbf{j}})) + \sum_{\mathbf{j}} \mu(C_{\mathbf{j}}) \sum_{t=0}^{T-1} S\Big(\frac{1}{\mu(C_{\mathbf{j}})} \int_{C_{\mathbf{j}}} d\mu \, \Theta^t(\mathrm{P}_\mathcal{F})\Big).$$

As the density matrices in the last term are convex combinations of pure states that remain close together, such a convex combination will be almost pure. The entropy of the resulting density matrix becomes then arbitrarily small when $\epsilon \to 0$. It suffices now to divide the inequality by T and to take the limit $T \to \infty$ in order to reach the conclusion of the theorem:

$$\mathsf{h}(\mathcal{F}) = \lim \frac{1}{T} S(\overline{\rho}_T) \leq \mathsf{h}^{\mathrm{KS}}. \quad \blacksquare$$

Let us define the dynamical entropy $\mathsf{h}_{\mathcal{B}}$ of (X, θ, μ) with respect to a Θ-invariant $*$-subalgebra of $\mathcal{L}^\infty(X, \mu)$ as the supremum of $\mathsf{h}(\mathcal{F})$ over all operational partitions $\mathcal{F}$ of unity in functions that belong to $\mathcal{B}$, we can then conclude that

THEOREM 3. *Both the continuous and measurable dynamical entropies of a classical dynamical system (X, θ, μ) coincide with the Kolmogorov-Sinai invariant of that system:*

$$\mathsf{h}^{\mathrm{KS}} = \mathsf{h}_{\mathcal{C}(X)} := \sup_{\mathcal{F} \subset \mathcal{C}(X)} \mathsf{h}(\mathcal{F}) = \mathsf{h}_{\mathcal{L}^\infty(X,\mu)} := \sup_{\mathcal{F} \subset \mathcal{L}^\infty(X,\mu)} \mathsf{h}(\mathcal{F}).$$

P r o o f. By the result of theorem 2 we must only show that h^{KS} is attained on continuous operational partitions of unity. Recalling the construction of the Kolmogorov-Sinai

invariant, we can find for any $\epsilon > 0$ a partition $\mathcal{C} = \{C_0, C_1, \ldots, C_{p-1}\}$ of X into disjoint measurable subsets such that the classical entropy

$$\lim_{T \to \infty} \frac{1}{T} \sum_{\mathbf{j}} \eta(\mu(C_{\mathbf{j}}))$$

exceeds $h^{\mathrm{KS}} - \epsilon$. As in the proof of theorem 2, $\mathbf{j}$ denotes a T-tuple $\{j_0, j_1, \ldots, j_T\}$ of integers between 0 and $p-1$ and $C_{\mathbf{j}}$ equals $\cap_t \theta^{-t}(C_{j_t})$.

The conditional entropy $\mathsf{H}[\mathcal{F} \mid \mathcal{G}]$ of a partition $\mathcal{F}$ with respect to a partition $\mathcal{G}$ is defined to be $\mathsf{H}[\mathcal{F} \circ \mathcal{G}] - \mathsf{H}[\mathcal{G}]$. In this expression $\mathsf{H}[\mathcal{F}]$ is the entropy of the partition $\mathcal{F} = \{f_0, f_1, \ldots, f_{q-1}\}$ i.e. the usual von Neumann entropy of the density matrix $[\mu(\overline{f_\ell} \, f_k)]_{k, \ell}$. By standard entropy inequalities [6] it is non-negative, subadditive in its first argument: $\mathsf{H}[\mathcal{F}_1 \circ \mathcal{F}_2 \mid \mathcal{G}] \leq \mathsf{H}[\mathcal{F}_1 \mid \mathcal{G}] + \mathsf{H}[\mathcal{F}_2 \mid \mathcal{G}]$ and monotonically decreasing in its second argument: $\mathsf{H}[\mathcal{F} \mid \mathcal{G}_1 \circ \mathcal{G}_2] \leq \mathsf{H}[\mathcal{F} \mid \mathcal{G}_1]$. We use these properties to obtain for any partition $\mathcal{F}$:

$$\mathsf{H}[\Theta^{T-1}(\mathcal{C}) \circ \cdots \circ \Theta(\mathcal{C}) \circ \mathcal{C}]$$
$$\leq \mathsf{H}[\Theta^{T-1}(\mathcal{F}) \circ \cdots \circ \Theta(\mathcal{F}) \circ \mathcal{F} \circ \Theta^{T-1}(\mathcal{C}) \circ \cdots \circ \Theta(\mathcal{C}) \circ \mathcal{C}]$$
$$\leq \mathsf{H}[\Theta^{T-1}(\mathcal{F}) \circ \cdots \circ \Theta(\mathcal{F}) \circ \mathcal{F}] + \sum_{j=0}^{T-1} \mathsf{H}[\Theta^j(\mathcal{C}) \mid \Theta^{T-1}(\mathcal{F}) \circ \cdots \circ \Theta(\mathcal{F}) \circ \mathcal{F}]$$
$$\leq \mathsf{H}[\Theta^{T-1}(\mathcal{F}) \circ \cdots \circ \Theta(\mathcal{F}) \circ \mathcal{F}] + \sum_{j=0}^{T-1} \mathsf{H}[\Theta^j(\mathcal{C}) \mid \Theta^j(\mathcal{F})]$$
$$= \mathsf{H}[\Theta^{T-1}(\mathcal{F}) \circ \cdots \circ \Theta(\mathcal{F}) \circ \mathcal{F}] + T \, \mathsf{H}[\mathcal{C} \mid \mathcal{F}].$$

Dividing by T and taking the limit for large T, we conclude that

$$h^{\mathrm{KS}} - \epsilon \leq \mathsf{h}[\mathcal{F}] + \mathsf{H}[\mathcal{C} \mid \mathcal{F}].$$

In order to finish the proof, we must show that we can find for any partition $\mathcal{C}$ of X into disjoint subsets a continuous operational partition $\mathcal{F}$ such that $\mathsf{H}[\mathcal{C} \mid \mathcal{F}] < \epsilon$. By the regularity of μ we can find for each of the C_j a compact K_j and an open O_j such that $K_j \subset C_j \subset O_j$ and such that the measures of $C_j \setminus K_j$ and $O_j \setminus C_j$ become arbitrarily small. Urysohn's lemma asserts the existence of continuous, non-negative, functions f_j that take the value 1 on K_j and vanish outside O_j. $\{\sqrt{f_0}, \sqrt{f_1}, \ldots, \sqrt{f_{p-1}}\}$ is, up to a continuous correction with small $\mathcal{L}^1$-norm, a continuous partition of unity, that coincides with $\mathcal{C}$, again up to a correction with small $\mathcal{L}^1$-norm. The continuity of the entropy ensures then that $\mathsf{H}[\mathcal{C} \mid \mathcal{F}]$ can be made less than ϵ. $\blacksquare$

In the proof of theorem 3, it appears that a restricted class of partitions can suffice to compute the Kolmogorov-Sinai invariant. In the classical computation, one has to study the proliferation of intersections of subsets of the phase space under the dynamics. This becomes quite complicated unless there exists a sufficiently large class of subsets of X whose boundaries transform nicely under the dynamics. Passing from the geometrical picture to a more analytic one allows us to use continuous functions which might allow for an easier implementation of the dynamics. Examples of such situations have been treated in [1]. The price to be paid is that we must now consider non-diagonal matrices. A special situation occurs when one is able to find a "Fourier" version of the dynamics in

the sense that one considers partitions in terms of complex-valued functions of constant modulus. A suitable choice of such partitions might then again lead to essentially diagonal density matrices if the, often wildly oscillating, phases of the products of different elements in the partition, average to zero. It should be noticed that partitions in functions of constant modulus preserve this property under composition and time evolution. The following result shows that such partitions indeed suffice to compute h^{KS}:

THEOREM 4. *The Kolmogorov-Sinai invariant h^{KS} of a dynamical system (X, θ, μ) can be computed as the supremum of the entropies $h[\mathcal{F}]$ of continuous operational partitions in functions of constant absolute value.*

P r o o f. Here too, it will be sufficient to show that we can approximate arbitrarily well the entropy of any partition $\mathcal{C} = \{C_0, C_1, \ldots, C_{q-1}\}$ of X in disjoint measurable sets by the entropy of a partition in measurable functions of constant absolute value. Passing from such partitions to continuous ones is possible with an approximation procedure similar to that in the proof of theorem 3.
Consider the partition $\mathcal{F} = \{f_0/\sqrt{q}, f_1/\sqrt{q}, \ldots, f_{q-1}/\sqrt{q}\}$ with f_j equal to $\exp(2\pi i j k/q)$ on C_k. A simple calculation shows that

$$[\mu(C_k)\, \delta_{k\,\ell}]_{k\,\ell} \sim [\mu(\overline{f_\ell}\, f_k)]_{k\,\ell},$$

where $\sim$ means unitary equivalence. This unitary equivalence is maintained under the time evolution. But this means that the entropy of the partition $\mathcal{F}$ equals that of the partition $\mathcal{C}$. ∎

It should be remarked that the partitions that appear in the proof of theorem 4 are constructed by slicing the phase space in disjoint pieces. In actual models, quite different, much smoother, partitions in functions of constant absolute value have proven useful. It is unclear for the moment under what conditions a single operational partition will actually produce the Kolmogorov-Sinai invariant.

References

[1]   R. A l i c k i, J. A n d r i e s, M. F a n n e s and P. T u y l s, *An algebraic approach to the Kolmogorov-Sinai entropy*, Rev. Math. Phys. 8 (1996), 167–184.

[2]   R. A l i c k i and M. F a n n e s, *Defining quantum dynamical entropy*, Lett. Math. Phys. 32 (1994), 75–82.

[3]   V.I. A r n o l d and A. A v e z, *Problèmes ergodiques de la mécanique classique*, Gauthier-Villars, Paris, 1967.

[4]   G. L i n d b l a d, *Non-Markovian stochastic processes and their entropy*, Commun. Math. Phys. 65 (1979), 281–294.

[5]   P. W a l t e r s, *An Introduction to Ergodic Theory*, Springer, Berlin, 1982.

[6]   A. W e h r l, *General properties of entropy*, Rev. Mod. Phys. 50 (1978), 221–260.