Journal of Statistical Physics, Vol. 24, No. 2, 1981

A Nonequilibrium Entropy for Dynamical Systems

S. Goldstein[1] **and O. Penrose**[2]

Received November 13, 1979

It is proposed to define entropy for nonequilibrium ensembles using a method of coarse graining which partitions phase space into sets which typically have zero measure. These are chosen by considering the totality of future possibilities for observation on the system. It is shown that this entropy is necessarily a nondecreasing function of the time t. There is no contradiction with the reversibility of the laws of motion because this method of coarse graining is asymmetric under time reversal. Under suitable conditions (which are stated explicitly) this entropy approaches the equilibrium entropy as $t \to +\infty$ and the fine-grained entropy as $t \to -\infty$. In particular, the conditions can always be satisfied if the system is a K-system, as in the Sinai billiard models. Some theorems are given which give information about whether it is possible to generate the partition used here for coarse graining from time translates of a finite partition, and at the same time elucidate the connection between our concept of entropy and the entropy invariant of Kolmogorov and Sinai.

KEY WORDS: Entropy; dynamical system; K-system; H-theorem; irreversibility.

1. A NONEQUILIBRIUM ENTROPY FOR DYNAMICAL SYSTEMS

The question of how to define entropy for nonequilibrium systems is of long standing [see the reviews by Wehrl[(1)] and Penrose[(2)]]. It is generally thought (though not universally—see Prigogine,[(3)] Misra[(4)]) that the right way to define a nonequilibrium entropy should involve some kind of coarse graining (see, for example, Tolman[(5)]). However, if the coarse graining is done in an arbitrary way, then the coarse-grained entropy is not necessarily a nondecreasing function of time. In this paper we propose a new type of

Research supported in part by NSF grants PHY78-03816 and PHY78-15920.

[1] Hill Center for Mathematical Sciences, Rutgers University, New Brunswick, New Jersey.
[2] Faculty of Mathematics, Open University, Milton Keynes, England.

0022-4715/81/0200-0325$03.00/0

coarse graining for dynamical systems which leads naturally to a definition of nonequilibrium entropy which does have the property of nondecrease with time, and under suitable conditions has the equally desirable property of tending to the equilibrium entropy as time proceeds.

The entropy defined here is quite distinct from the Kolmogorov–Sinai entropy invariant used in ergodic theory. Our entropy is a time-dependent property of nonequilibrium measures, whereas that of Kolmogorov and Sinai is independent of time and depends only on the measure-theoretic structure of the dynamical system. There is, however, a connection between the two entropy concepts: the Kolmogorov–Sinai invariant is related to the rate at which the entropy we define here can increase with time.

2. THE OBSERVATIONAL MODEL

Our method of coarse graining depends on a specific model of the observational process, which we now describe.

We describe the microscopic states of the system by means of "phase points" ω and denote the phase space consisting of all possible phase-space points by Ω. We shall say that two phase-space points ω_1 and ω_2 are *observationally equivalent* if two systems, one having the phase point ω_1 and one the phase point ω_2, will, if subjected to precisely the same experimental procedures, give precisely the same observable behavior throughout the future. We make the assumption (which could be challenged, as in the theory of "fuzzy observables,"[(6–8)]) that if ω_1 is observationally equivalent to ω_2 and ω_2 is observationally equivalent to ω_3, then ω_1 is observationally equivalent to ω_3; in this case observational equivalence is an equivalence relation in the mathematical sense. We can then split phase space into a family or *partition* Π of nonoverlapping sets, such that two phase points belong to the same set in Π if and only if they are observationally equivalent.

Since its definition is not symmetric between past and future, the partition Π itself is not symmetric between past and future. We can characterize this asymmetry using the family of time evolution transformations ϕ_t defined for all real t by the condition that a system whose phase point is ω at time 0 will have phase point $\phi_t\omega$ at time t. For (isolated) dynamical systems these transformations are invertible and have the group property $\phi_s\phi_t = \phi_{s+t}$ for all real s, t.

Suppose now that ω_1 and ω_2 are any two phase points belonging to the same set α in the partition Π. Since ω_1 and ω_2 cannot be distinguished by any observational procedure starting at time $t = 0$, then *a fortiori* they cannot be distinguished by any observational procedure starting at a later time t (with $t > 0$); hence the points $\phi_t\omega_1$ and $\phi_t\omega_2$ into which the time evolution carries them are observationally equivalent and therefore belong to the same set, call it β, in the partition Π. So if α is any one of the sets belonging to Π, then the set $\phi_t\alpha$, consisting of all the images under ϕ_t of points in α, is a subset of some set β in Π.

Since ϕ_t is invertible, every phase point in Ω belongs to one of the sets $\phi_t\alpha$ ($\alpha \in \Pi$); these sets therefore constitute a partition of Ω, which we may denote by $\phi_t\Pi$; in other words, the partition $\phi_t\Pi$ is a refinement of the partition Π. This relation between Π and $\phi_t\Pi$ is the asymmetric property we require, since it holds only for nonnegative t.

Since the definition of entropy involves integration, we want to replace the above condition by one involving measurable sets. We may safely assume the sets constituting the partition Π to be measurable and consider the family $\mathcal{Q}$ consisting of all measurable sets that are unions of sets from the family Π. If we also include the empty set $\varnothing$ in $\mathcal{Q}$, then $\mathcal{Q}$ is closed under the operations of forming complements and countable unions, and is called the σ-algebra generated by the partition Π. As we have just seen, every set in Π is the image under ϕ_t of some union of sets in Π; every set in $\mathcal{Q}$ is therefore the image of some union of sets in Π. Moreover, since ϕ_t is measure-preserving and invertible, this union of sets in Π is measurable and therefore belongs to $\mathcal{Q}$. Thus every set in $\mathcal{Q}$ is the image under ϕ_t of some set in $\mathcal{Q}$; that is to say, the algebra $\phi_t\mathcal{Q}$, consisting of the images under ϕ_t of all the sets belonging to $\mathcal{Q}$, includes (among others) all the sets of $\mathcal{Q}$ itself:

$$\phi_t\mathcal{Q} \supset \mathcal{Q} \qquad (t > 0) \tag{2.1}$$

The σ-algebra $\mathcal{Q}$ is our mathematical representation of the observational possibilities available at any given instant, and the condition (2.1) represents the loss of observational or information-gathering possibilities as time proceeds. It will be the source of our results about the increase of entropy with time.

A particular case of (2.1), which we shall use later, arises if we further specialize the model of observation by making the following two natural assumptions:

i. Observations are possible only at an equally spaced set of instants, which we denote by $t = \ldots, -2, -1, 0, 1, 2, 3, \ldots$.
ii. The observation made at any instant has only a finite set of possible outcomes.

Let us define P to be the partition of phase space into sets of points that are indistinguishable by an observation made at time 0. Two phase points ω_1 and ω_2 are then observationally equivalent if and only if the time translates $\phi_t\omega_1$ and $\phi_t\omega_2$ lie, for every nonnegative integer t, in the same set from the partition P. In other words, ω_1 and ω_2 must lie in the same set from the partition $\phi_{-t}P$. Hence the sets forming the partition Π are intersections of sets from the partitions $P, \phi_{-1}P, \phi_{-2}P, \ldots$. The σ-algebra $\mathcal{Q}$ therefore includes all the sets belonging to $P, \phi_{-1}P, \phi_{-2}P, \ldots$, together with sets formed by taking countable intersections of these sets. Indeed, it can be defined as the smallest σ-algebra which contains all the partitions $P, \phi_{-1}P, \phi_{-2}P, \ldots$, written

$$\mathcal{Q} = \bigvee_{t=0}^{\infty} \phi_{-t}P \tag{2.2}$$

3. DEFINITION OF ENTROPY

Let μ be some finite invariant measure (not necessarily normalized) on the phase space Ω, and let ν be any measure, absolutely continuous with respect to μ, which is normalized but not necessarily invariant. We may think of ν as defined by a normalized density on the phase space Ω, given by the Radon–Nikodym derivative

$$\rho = d\nu / d\mu$$

We shall assume it to have a finite fine-grained entropy:

$$h(\nu) = -k \int \eta(\rho)\, d\mu > -\infty \tag{3.1}$$

where

$$\eta(x) = \begin{cases} x \log x & (x > 0) \\ 0 & (x = 0) \end{cases} \tag{3.2}$$

and k is Boltzmann's constant, a positive number.

For the problem of predicting the observable future behavior of the system, the fine-grained density ρ is too detailed; it contains information which does not affect future observational possibilities. This unnecessary information can be removed, without affecting the probabilities of observable future events, if we replace ρ by a different probability density $\rho_{\mathcal{Q}}$ obtained by averaging ρ over each equivalence class of observable states. The mathematical formulation of this "averaging" is provided by the conditional expectation of ρ with respect to the algebra $\mathcal{Q}$,[3]

$$\rho_{\mathcal{Q}} = E(\rho \mid \mathcal{Q}) \tag{3.3}$$

[3] This is the conditional expectation for the measure $\mu / \mu(\Omega)$.

We therefore adopt, as our definition of coarse-grained entropy, the formula

$$h(\nu, \mathcal{Q}) = -k \int \eta(\rho_{\mathcal{Q}})\, d\mu \tag{3.4}$$

Put differently, the entropy should depend on ν and μ regarded as measures on the σ-algebra $\mathcal{Q}$ of observational equivalence rather than the full σ-algebra. So regarded, the derivative of ν with respect to μ is just (3.3). Note also that $h(\nu, \mathcal{Q})$ is well defined for *any* sub-σ-algebra $\mathcal{Q}$.

4. TIME EVOLUTION

The time evolution of the measure ν consists of the family of measures $\{\nu_t\}$ defined (formally) by

$$\nu_t(\phi_t \omega) = \nu(\omega) \tag{4.1}$$

This implies, for any set A in phase space, that

$$\nu_t(A) = \nu(\phi_{-t} A) \tag{4.2}$$

and hence, since the measure μ used in defining the entropy is invariant, that

$$h(\nu_t, \mathcal{Q}) = h(\nu, \phi_{-t}\mathcal{Q}) \tag{4.3}$$

The nondecrease property of the entropy now follows from the following result.

Theorem 1. $h(\nu, \mathcal{B})$ is a nonincreasing function of $\mathcal{B}$. That is, if $\mathcal{B}$ and $\mathcal{C}$ are any two σ-algebras of measurable sets on Ω, satisfying $\mathcal{B} \subset \mathcal{C}$, then

$$h(\nu, \mathcal{B}) \geqslant h(\nu, \mathcal{C}) \tag{4.4}$$

Proof. The basis of our proof is Jensen's inequality,[9] which implies, since the function η is convex, that

$$\eta(E\rho_{\mathcal{C}} \mid \mathcal{B})) \leqslant E(\eta(\rho_{\mathcal{C}}) \mid \mathcal{B}) \qquad \text{a.e.} \tag{4.5}$$

Since $\mathcal{B} \subset \mathcal{C}$, we have, by a well-known property of conditional expectations (Ref. 9, Theorem 10.2)

$$E(E(\rho \mid \mathcal{C}) \mid \mathcal{B}) = E(\rho \mid \mathcal{B}) \qquad \text{a.e.} \tag{4.6}$$

In our notation, defined in (3.3), this becomes

$$E(\rho_{\mathcal{C}} \mid \mathcal{B}) = \rho_{\mathcal{B}} \qquad \text{a.e.} \tag{4.7}$$

The definition of conditional expectation implies that

$$\int_\Omega E(\eta(\rho_{\mathcal{C}}) \mid \mathcal{B})\, d\mu = \int_\Omega \eta(\rho_{\mathcal{C}})\, d\mu \tag{4.8}$$

Combining (4.5) and (4.7), integrating both sides of the resulting inequality over Ω, and using (4.8), we obtain

$$\int_{\Omega}\eta(\rho_{\mathfrak{B}})\,d\mu \leqslant \int_{\Omega}\eta(\rho_{\mathcal{C}})\,d\mu \tag{4.9}$$

By the definition (3.4) of h, this last result is equivalent to the required inequality (4.4). ■

Corollary 1.1. The entropy $h(\nu,\mathcal{Q})$ is bounded below by the fine-grained entropy $h(\nu)$ and above by the "equilibrium" entropy $k\log\mu(\Omega)$.

Proof. Theorem 1 implies

$$h(\nu,\{\varnothing,\Omega\}) \geqslant h(\nu,\mathcal{Q}) \geqslant h(\nu,\mathcal{F})$$

where $\mathcal{F}$ is the algebra of all measurable sets on Ω. Since

$$h(\nu,\{\varnothing,\Omega\}) = k\log\mu(\Omega) \qquad \text{and} \quad h(\nu,\mathcal{F}) = h(\nu)$$

the result follows. ■

Corollary 1.2 (An "H-Theorem"). The entropy $h(\nu_t,\mathcal{Q})$ is a nondecreasing function of t.

Proof. If $t > s$, then (2.1) implies $\phi_{t-s}\mathcal{Q} \supset \mathcal{Q}$, so that

$$\phi_{-s}\mathcal{Q} \supset \phi_{-t}\mathcal{Q}$$

It follows, by (4.3) and Theorem 1, that

$$\begin{aligned} h(\nu_t,\mathcal{Q}) &= h(\nu,\phi_{-t}\mathcal{Q}) \\ &\geqslant h(\nu,\sigma_{-s}\mathcal{Q}) \\ &= h(\nu_s,\mathcal{Q}) \qquad (t > s) \quad ■ \end{aligned}$$

5. CONTINUITY PROPERTIES OF THE ENTROPY

Theorem 1 shows that the entropy is bounded below by its fine-grained value $h(\nu)$ defined in (3.1), and above by

$$h(\nu,\{\varnothing,\Omega\}) = k\log\mu(\Omega) \tag{5.1}$$

The nondecrease property therefore implies that $h(\nu_t,\mathcal{Q})$ must approach limits as $t\to+\infty$ and $t\to-\infty$. We should like to be able to show that these limits are equal to the corresponding upper and lower bounds.

This question can be investigated using the following continuity property.

Theorem 2. If $\{\mathfrak{B}_t\}$ is an increasing or decreasing family of σ-algebras, with t either continuous or discrete, and if $\mathfrak{B}_\infty$ is its limit in the sense that $\mathfrak{B}_t \uparrow \mathfrak{B}_\infty$ or $\mathfrak{B}_t \downarrow \mathfrak{B}_\infty$, then

$$\lim_{t\to\infty} h(\nu, \mathfrak{B}_t) = h(\nu, \mathfrak{B}_\infty) \tag{5.2}$$

Proof. By the definition (3.4) the result we wish to prove is equivalent to

$$\lim_{t\to\infty} \int_\Omega f_t \, d\mu = \int_\Omega f_\infty \, d\mu \tag{5.3}$$

where

$$f_t = \eta(\rho_t) \tag{5.4}$$

with

$$\rho_t = E(\rho \mid \mathfrak{B}_t) \tag{5.5}$$

and f_∞ is defined analogously.

The martingale convergence theorem (Refs. 10, 11; see also Ref. 9, pp. 116 and 121) tells us that

$$\lim_{t\to\infty} f_t(\omega) = f_\infty(\omega) \qquad \text{a.e.} \tag{5.6}$$

and so if the density $\rho(\omega)$ is bounded, the result (5.3) follows from Lebesgue's dominated convergence theorem. In general, however, $\rho(\omega)$ need not be bounded, nor need the family $\{f_t\}$ be bounded by an integrable function, and so a more complicated proof is necessary. Our proof depends on the following result.

Lemma 2.1. The family of functions f_t is *uniformly integrable*; that is to say, the function λ defined by

$$\lambda(x) = \sup_t \int_\Delta |f_t| \, d\mu \tag{5.7}$$

where

$$\Delta = \Delta(x,t) = \{\omega: \quad |f_t(\omega)| > x\} \tag{5.8}$$

satisfies

$$\lim_{x\to\infty} \lambda(x) = 0 \tag{5.9}$$

Proof of Lemma 2.1. Since we are interested only in large x, we need only consider values larger than $1/e$. Then, since the function η is bounded below by $-1/e$, only positive values of f_t contribute to the integral in (5.7), so that

$$\int_\Delta |f_t| \, d\mu = \int_\Delta f_t \, d\mu \qquad (x > 1/e) \tag{5.10}$$

Jensen's inequality, with (5.4) and (5.5), gives

$$\int_\Delta f_t\,d\mu = \int_\Delta \eta(E(\rho\,|\,\mathcal{B}_t))\,d\mu$$
$$\leqslant \int_\Delta E(\eta(\rho)\,|\,\mathcal{B}_t)\,d\mu$$
$$= \int_\Delta \eta(\rho)\,d\mu \tag{5.11}$$

since Δ belongs to $\mathcal{B}_t$.

The set Δ may be divided into two parts: Δ_1, in which $\eta(\rho) \geqslant x^{1/2}$ and Δ_2, in which $\eta(\rho) < x^{1/2}$. The contribution of Δ_1 to the last integral in (5.11) is bounded above by

$$\lambda_1(x) = \int_{\eta(\rho) \geqslant x^{1/2}} \eta(\rho)\,d\mu \tag{5.12}$$

and because of (3.1) we have

$$\lim_{x\to\infty} \lambda_1(x) = 0 \tag{5.13}$$

Throughout Δ_2 we have $\eta(\rho) < x^{1/2}$ and $f_t \geqslant x$, so that $\eta(\rho) \leqslant f_t/x^{1/2}$, and the contribution of Δ_2 to the integral is bounded above by

$$x^{-1/2}\int_{\Delta_2} f_t\,d\mu \leqslant x^{-1/2}\int_\Delta f_t\,d\mu \qquad (x > 1/e) \tag{5.14}$$

Using the estimates (5.12) and (5.14) in (5.11) and rearranging, we obtain

$$(1 - x^{-1/2})\int_\Delta f_t\,d\mu \leqslant \lambda_1(x) \qquad (x > 1/e) \tag{5.15}$$

Combining this with (5.10) and the definition (5.7) of $\lambda(x)$, we find that

$$0 \leqslant \lambda(x) \leqslant \lambda_1(x)/(1 - x^{-1/2}) \qquad (x > 1) \tag{5.16}$$

Using (5.13), we complete the proof of (5.9) and hence of the lemma. ■

Returning now to the proof of Theorem 2, let us define

$$f_t^{(x)}(\omega) = \begin{cases} f_t(\omega) & \text{if} \quad |f_t(\omega)| < x \\ 0 & \text{if} \quad |f_t(\omega)| \geqslant x \end{cases} \tag{5.17}$$

and $f_\infty^{(x)}$ similarly. Equation (5.7) of Lemma 2.1 then implies, since

$$\int_\Delta |f_t|\,d\mu = \int_\Omega |f_t - f_t^{(x)}|\,d\mu$$

that

$$\int_\Omega f_t\,d\mu \lessgtr \int_\Omega f_t^{(x)}\,d\mu \pm \lambda(x) \tag{5.18}$$

The integrability of the function f_t is guaranteed by the formula (5.18) itself, or alternatively by the fact that

$$k\int f_t\,d\mu = -h(\nu, \mathcal{B}_t) \leqslant -h(\nu)$$

To prove Theorem 2, we now take the limits $t \to \infty$ and then $x \to \infty$ on both sides of (5.18). For the limit $t \to \infty$ we use the martingale theorem (5.6) which implies

$$\lim_{t\to\infty} f_t^{(x)}(\omega) = f_\infty^{(x)}(\omega) \qquad \text{a.e.} \tag{5.19}$$

provided x is such that

$$\mu\{\omega:\ \ f_\infty(\omega) = x\} = 0 \tag{5.20}$$

The set of values of x for which (5.20) is false will be denoted by F; it is a countable set of real numbers. Applying Lebesgue's dominated convergence theorem to (5.19), we obtain

$$\lim_{t\to\infty} \int_\Omega f_t^{(x)} = \int_\Omega f_\infty^{(x)}\, d\mu \tag{5.21}$$

and so the $t \to \infty$ limit of (5.18) can be written

$$\lim_{t\to\infty} \left\{ {\sup \atop \inf} \right\} \int_\Omega f_t\, d\mu \lessgtr \int_\Omega f_\infty^{(x)}\, d\mu \pm \lambda(x) \tag{5.22}$$

Finally, we take the limit $x \to \infty$, avoiding values of x in the countable set F. Since $\lambda(x) \to 0$ by Eq. (5.9) of Lemma 2.1, the result is simply Eq. (5.3). ■

From this theorem we can obtain sufficient conditions for $h(\nu_t, \mathscr{Q})$ to approach, in the limits $t \to +\infty$ and $t \to -\infty$, the upper and lower bounds mentioned in Corollary 1.1.

Corollary 2.2. (i) If $\phi_s \mathscr{Q} \uparrow \hat{1}$ as $s \to \infty$, where $\hat{1}$ denotes the algebra of all measurable sets, then $h(\nu_t, \mathscr{Q}) \to h(\nu)$ as $t \to -\infty$.

(ii) If $\phi_s \mathscr{Q} \downarrow \hat{0}$ as $s \to -\infty$, where $\hat{0}$ is the algebra of all sets of measure 0 or $\mu(\Omega)$, then $h(\nu_t, \mathscr{Q}) \to k \log \mu(\Omega)$ as $t \to +\infty$ [$k \log \mu(\Omega)$ is the thermodynamic equilibrium entropy].

(iii) If the dynamical system (ϕ, μ, Ω) is a K-system, then there is a sub-σ-algebra $\mathscr{Q}$ such that $h(\nu_t, \mathscr{Q})$ is nondecreasing, approaches the fine-grained entropy $h(\nu)$ as $t \to -\infty$, and approaches the thermodynamic entropy $k \log \mu(\Omega)$ as $t \to +\infty$. Moreover, every sub-σ-algebra $\mathscr{Q}$ of the form (2.2) for some finite partition P will satisfy $\phi_s \mathscr{Q} \downarrow \hat{0}$ in this case.

Proof. For (i) take $\mathscr{B}_s = \phi_{-s} \mathscr{Q}$ and $\mathscr{B}_\infty = \hat{1}$ in Theorem 2. For (ii) take $\mathscr{B}_s = \phi_s \mathscr{Q}$ and $\mathscr{B}_\infty = \hat{0}$. For (iii) use the definition of a K-system, which is simply that there should exist a subalgebra $\mathscr{Q}$ with the nonincrease property (2.1) and the two properties required in parts (i) and (ii). The "moreover" states the fact that for K-systems finite partitions have trivial tails (Ref. 11, Chap. 7). ■

Sinai[12] has argued that the hard-sphere gas is a K-system, using an increasing subalgebra $\mathscr{Q}$ defined in terms of the sequence of collisions that takes place between the various spheres. If this is so, then Eq. (3.4) defines a nondecreasing entropy for such a system in terms of the probability measure for sequences of future collisions.

Two further corollaries, though not in the mainstream of this work, are included for completeness.

Corollary 2.3. The following alternative definition of $h(\nu, \mathcal{Q})$ is equivalent to our definition (3.4):

$$h(\nu, \mathcal{Q}) = -k \inf_{\{A_i\} \in \mathcal{Q}} \sum_i \mu(A_i)\eta[\nu(A_i)/\mu(A_i)] \tag{5.23}$$

where the "inf" is taken over all finite partitions of Ω whose elements are in $\mathcal{Q}$.

Proof. Since every sub-σ-algebra of a Lebesgue space is countably generated (mod 0), there is an increasing sequence of finite partitions $P^{(1)}, P^{(2)}, \ldots$ such that $P^{(n)} \uparrow \mathcal{Q}$ as $n \to \infty$. For each of these finite partitions we have from the definition (3.4)[4]

$$h(\nu, \mathcal{P}^{(n)}) = -k \sum_{A \in P^{(n)}} \mu(A)\eta[\nu(A)/\mu(A)] \tag{5.24}$$

and, by Theorem 2,

$$\lim_{n\to\infty} h(\nu, \mathcal{P}^{(n)}) = h(\nu, \mathcal{Q}) \tag{5.25}$$

so that the right-hand side of (5.23) cannot exceed $h(\nu, \mathcal{Q})$. But, on the other hand, Theorem 1 shows that for *any* finite partition P with elements in $\mathcal{Q}$ we have

$$h(\nu, \mathcal{P}) \geqslant h(\nu, \mathcal{Q})$$

so that the right-hand side of (5.23) cannot be less than $h(\nu, \mathcal{Q})$. Thus the two sides of (5.23) are equal. ■

Corollary 2.4. Under the condition (ii) of Corollary 2.2, namely,

$$\phi_s \mathcal{Q} \downarrow \hat{0} \qquad \text{as} \quad s \to -\infty$$

the measures whose densities are $(d\nu_t/d\mu)_{\mathcal{Q}}$ converge as $t \to \infty$ to the equilibrium measure with constant density $\rho_\infty = 1/\mu(\Omega)$ not only in entropy, as implied by Corollary 2.2, but also in L^1.

Proof. We want to show that

$$\lim_{t\to\infty} \int_\Omega \left| E\left(\frac{d\nu_t}{d\mu} \,|\, \mathcal{Q} \right)(\omega) - \rho_\infty(\omega) \right| d\mu(\omega) = 0$$

Since ϕ_t preserves measure, this is equivalent to

$$\lim_{t\to\infty} \int_\Omega \left| E\left(\frac{d\nu_t}{d\mu} \,|\, \mathcal{Q} \right)[\phi_t(\omega)] - \rho_\infty[\phi_t(\omega)] \right| d\mu(\omega) = 0$$

[4] We write $\mathcal{P}$ for the sub-σ-algebra associated with the finite partition P.

which by (4.1) is the same as

$$\lim_{t\to\infty}\int_\Omega\left|E\left(\frac{d\nu}{d\mu}\,|\,\phi_{-t}\mathcal{C}\right)(\omega)-\rho_\infty(\omega)\right|d\mu(\omega)=0$$

This in turn can be written

$$\lim_{t\to\infty}\int_\Omega \bar{f}_t\,d\mu=\int_\Omega \bar{f}_\infty\,d\mu$$

where

$$\bar{f}_t=\bar{\eta}(\rho_t)\quad\text{and}\quad \bar{f}_\infty=\bar{\eta}(\rho_\infty)$$

with

$$\bar{\eta}(x)=|x-\rho_\infty|\qquad(x\in\mathbb{R})$$

and

$$\rho_t=E(\rho\,|\,\mathcal{B}_t)$$

with

$$\mathcal{B}_t=\phi_{-t}\mathcal{C}$$

Since $\bar{\eta}$ is a convex function, bounded below like η, the proof of Theorem 2 applies here too. ■

An alternative proof is to use the L^1 martingale convergence theorem.

6. THE BAKER'S TRANSFORMATION

To illustrate some of the results in this paper, we consider two examples. The first is the baker's transformation,[13] in which the phase space is the unit square $\{(p,q)\colon\ 0\leqslant p<1,\ 0\leqslant q<1\}$ and the time evolution is defined by

$$\phi_1(p,q)=\begin{cases}(2p,\tfrac{1}{2}q) & \text{if}\ \ p<\tfrac{1}{2}\\ (2p-1,\tfrac{1}{2}q+\tfrac{1}{2}) & \text{if}\ \ p\geqslant\tfrac{1}{2}\end{cases}\tag{6.1}$$

together with the obvious rule

$$\phi_t=(\phi_1)^t\qquad(t\in\mathbb{Z})$$

A suitable algebra $\mathcal{C}$ can be constructed by assuming that the only observation possible is to observe whether or not $p<\frac{1}{2}$, and using the recipe (2.2). The partition P divides phase space into two rectangles: in one $0\leqslant p<\frac{1}{2}$ and in the other $\frac{1}{2}\leqslant p<1$. The formula (2.2) then tells us that $\mathcal{C}$ corresponds to the partition of phase space into lines $p=\text{const}$.

For any absolutely continuous measure ν, the coarse-grained probability density $\rho_{\mathcal{C}}$ corresponding to the algebra $\mathcal{C}$ is given, in accordance with

(3.3), by

$$\rho_{\mathcal{C}}(p,q) = \sigma(p) \tag{6.2}$$

where

$$\sigma(p) = \int_0^1 \rho(p,q)\,dq$$

with ρ defined by $d\nu = \rho(p,q)\,dp\,dq$. Its time evolution law is given by[14]

$$\sigma_{t+1}(p) = \tfrac{1}{2}\left[\sigma_t(\tfrac{1}{2}p) + \sigma_t(\tfrac{1}{2}p + \tfrac{1}{2})\right] \tag{6.3}$$

The entropy associated with the measure ν is, by (3.4),

$$h(\nu,\mathcal{C}) = -k\int_0^1 \sigma(p)\log\sigma(p)\,dp \tag{6.4}$$

Since the function $\eta(x) = x\log x$ is convex, it follows by (6.3) and Jensen's inequality that

$$\begin{aligned} h(\nu_{t+1},\mathcal{C}) &= -k\int_0^1 \eta(\sigma_{t+1}(p))\,dp \\ &\geqslant -k\int_0^1 \left[\tfrac{1}{2}\eta(\sigma_t(\tfrac{1}{2}p)) + \tfrac{1}{2}\eta(\sigma_t(\tfrac{1}{2}p + \tfrac{1}{2}))\right] dp \\ &= -k\left[\int_0^{1/2} \eta(\sigma_t(p'))\,dp' + \int_{1/2}^1 \eta(\sigma_t(p'))\,dp'\right] \\ &= h(\nu_t,\mathcal{C}) \end{aligned} \tag{6.5}$$

which confirms the "H-theorem" of Corollary 1.2.

According to Corollary 1.1, $h(\nu,\mathcal{C})$ has the upper and lower bounds

$$h(\nu) \leqslant h(\nu,\mathcal{C}) \leqslant k\log 1 = 0 \tag{6.6}$$

where

$$h(\nu) = -k\int_0^1\int_0^1 \rho(p,q)\log\rho(p,q)\,dp\,dq$$

is the fine-grained entropy. Moreover, since the baker's transformation is a K-system,[15] Corollary 2.2 tells us that $h(\nu_t,\ \mathcal{C})$ approaches its upper bound 0 as $t\to\infty$ and its (negative) lower bound $h(\nu)$ as $t\to-\infty$.

7. MARKOV CHAINS

The second of our two examples is the Markov chain. Let Γ be a space on which is defined a σ-algebra $\mathcal{F}$ of sets, and denote the transition kernel by

$$K(\gamma,F) \qquad (\gamma\in\Gamma,\quad F\in\mathcal{F})$$

This means that if the probability distribution over Γ at time t is given by a measure $\nu_{\Gamma,t}$ the probability distribution at time $t+1$ is given by the measure $\nu_{\Gamma,t+1} = \nu_{\Gamma,t}K$ defined by

$$\nu_{\Gamma,t+1}(F) = \int_\Gamma d\nu_{\Gamma,t}(\gamma)K(\gamma, F) \tag{7.1}$$

Let μ_Γ be an equilibrium measure on Γ, that is, one satisfying $\mu_\Gamma K = \mu_\Gamma$, and suppose that $\nu_{\Gamma,t}$ is absolutely continuous with respect to μ_Γ. Then one would expect to be able to show, using Jensen's inequality, that the entropy defined by

$$h_\Gamma(\nu_\Gamma) = -k\int_\Gamma d\mu_\Gamma \eta\left(\frac{d\nu_\Gamma}{d\mu_\Gamma}\right) \tag{7.2}$$

satisfies an H-theorem:

$$h_\Gamma(\nu_{\Gamma,t+1}) \geqslant h_\Gamma(\nu_{\Gamma,t}) \tag{7.3}$$

The original H-theorem of Pauli[16] is a special case of this formula and there have been various generalizations of Pauli's result since—for example, that of Lindblad.[17]

To relate these properties of Markov chains to the results of the present paper we consider an abstract dynamical system associated with the Markov chain. Its phase space Ω is the set of all doubly infinite sequences

$$\omega = \{\ldots, \omega_{-1}, \omega_0, \omega_1, \ldots\} \tag{7.4}$$

with

$$\omega_n \in \Gamma \qquad \text{for all integers } n$$

The time evolution $\{\phi_t\}_{t\in\mathbb{Z}}$ on Ω is defined by

$$(\phi_t\omega)_n = \omega_{n+t} \tag{7.5}$$

We may define a σ-algebra over Ω by

$$\bigvee_{n=-\infty}^{\infty} \mathcal{F}_n \tag{7.6}$$

where $\mathcal{F}_n$ is a sub-σ-algebra over Ω consisting of sets of the form

$$\{\omega: \quad \omega_n \in F\} \qquad \text{with } F \in \mathcal{F} \tag{7.7}$$

For our sub-σ-algebra $\mathcal{Q}$ we choose, by analogy with (2.2),

$$\mathcal{Q} = \bigvee_{n\geqslant 0} \mathcal{F}_n \tag{7.8}$$

Its time translates are given by

$$\phi_{-t}\mathcal{Q} = \bigvee_{n\geqslant t} \mathcal{F}_n \tag{7.9}$$

so that our basic "loss of information" condition (2.1) is satisfied.

Associated with the equilibrium measure μ_Γ over the σ-algebra $\mathfrak{F}$ in Γ there is a stationary measure μ over the σ-algebra $\bigvee_{n=-\infty}^{\infty}\mathfrak{F}_n$ in Ω defined by

$$\begin{aligned} &\mu\{\omega:\quad \omega_n \in F\} = \mu_\Gamma(F) \qquad (n \in \mathbb{Z};\quad F \in \mathfrak{F}) \\ &\mu\{\omega:\quad \omega_n \in F_n \quad \text{and} \quad \omega_{n+1} \in F_{n+1}\} \\ &\qquad = \int_{F_n} d\mu_\Gamma(\omega_n)\, K(\omega_n, F_{n+1}) \qquad (F_n, F_{n+1} \in \mathfrak{F}) \end{aligned} \tag{7.10}$$

etc.

Moreover, with any family of nonequilibrium measures $\{\nu_{\Gamma,t}\}_{t\geqslant 0}$ in Γ satisfying (7.1), there is associated a nonstationary measure ν over the subalgebra $\mathcal{Q}$ in Γ, defined analogously:

$$\begin{aligned} &\nu\{\omega:\quad \omega_n \in F\} = \nu_{\Gamma,n}(F) \qquad (n \geqslant 0,\quad F e \mathfrak{F}) \\ &\nu\{\omega:\quad \omega_n \in F_n \quad \text{and} \quad \omega_{n+1} \in F_{n+1}\} = \int_{F_n} d\nu_{\Gamma,n}(\omega_n)\, K(\omega_n, F_{n+1}) \end{aligned} \tag{7.11}$$

etc.

If $\nu_{\Gamma,0}$ is absolutely continuous with respect to μ_Γ, then ν is absolutely continuous with respect to $\mu_{\mathcal{Q}}$, the restriction of μ to $\mathcal{Q}$, and the Radon–Nikodym derivatives are essentially the same, in the sense that $d\nu/d\mu_{\mathcal{Q}}$ is measurable in $\mathfrak{F}_0$ and is isomorphic in $\mathfrak{F}_0$ to the measure $d\nu_{\Gamma,0}/d\mu_\Gamma$ in $\mathfrak{F}$. We may write this relation as

$$\left(\frac{d\nu}{d\mu}\right)_{\mathfrak{F}_0} \sim \frac{d\nu_{\Gamma,0}}{d\mu_\Gamma}$$

and it has the consequence that

$$h(\nu, \mathcal{Q}) = h_\Gamma(\nu_{\Gamma,0}) \tag{7.12}$$

In a similar way, we can show that

$$h(\nu, \phi_{-t}\mathcal{Q}) = h_\Gamma(\nu_{\Gamma,t}) \tag{7.13}$$

and so the Markov chain H-theorem (7.3) is seen to be a corollary of our more general H-theorem given in Corollary 1.2.

If in addition the Markov chain has trivial tail, e.g., if it has a finite state space and is irreducible and aperiodic, then the algebra $\mathcal{Q}$ satisfies the "trivial tail" condition (ii) of Corollary 2.2: $\phi_s\mathcal{Q}\downarrow\hat{0}$ as $s \to -\infty$. In that case Corollaries 2.2 and 2.4 tell us that, as $t \to \infty$, $d\nu_{\Gamma,t}/d\mu_\Gamma$ converges in entropy and also in L^1 to a unique equilibrium phase-space density, which is a constant.

8. RELATION TO THE KOLMOGOROV–SINAI ENTROPY

In the two examples just considered, it was easy to find a σ-algebra $\mathcal{Q}$ satisfying our fundamental condition $\phi_t\mathcal{Q} \supset \mathcal{Q}$ in a sufficiently nontrivial way to ensure that $h(\nu_t, \mathcal{Q})$ is not constant. For more general dynamical systems it is not so obvious that this can be done, let alone whether $\mathscr{Q}$ can be constructed from a finite partition using the recipe (2.3). The following theorems give some information about this question, and at the same time elucidate the connection between our concept of entropy and the entropy invariant of Kolmogorov and Sinai.

First we need some notation. Let $\phi \equiv \phi_1$, and let $H = H(\phi)$ denote the Kolmogorov–Sinai (KS) entropy invariant. For any finite partition P let $\mathscr{Q}_P$ denote the σ-algebra of the form (2.2) generated by P.

For $\mathscr{Q} \subset \phi\mathscr{Q}$ the *entropy increase in state* ν (for ϕ given $\mathscr{Q}$) is defined by

$$\Delta^{\phi}(\nu, \mathscr{Q}) = h(\nu_1, \mathscr{Q}) - h(\nu, \mathscr{Q}) = h(\nu, \phi^{-1}\mathscr{Q}) - h(\nu, \mathscr{Q})$$

and the *entropy increase for* ϕ *given* $\mathscr{Q}$ is defined by

$$\Delta^{\phi}(\mathscr{Q}) = \sup_{\nu} \Delta^{\phi}(\nu, \mathscr{Q})$$

where the "sup" is over all $\nu \ll \mu$ with $h(\nu) > -\infty$. Suppose ϕ is ergodic with $H < \infty$. Then we define the *entropy increase for* ϕ by

$$\Delta^{\phi} = \inf \Delta^{\phi}(\mathscr{Q})$$

where the "inf" is over all σ-algebras $\mathscr{Q} = \mathscr{Q}_P$ "of full entropy," i.e., for which $H(P, \phi)$—the KS entropy of ϕ given P—equals H (Ref. 11, Chap. 5). Finally, the *asymptotic rate of entropy increase for* ϕ is given by

$$\Delta^{\phi}_{\mathrm{as}} = \lim_{t \to \infty} (1/t) \Delta^{\phi_t} \qquad (t \in \mathbb{Z})$$

Theorem 3. (a) For $\mathscr{Q} \subset \phi\mathscr{Q}$, $\Delta^{\phi}(\mathscr{Q}) > 0 \Leftrightarrow H(\mathscr{Q} \| \phi^{-1}\mathscr{Q}) > 0$.[5] In particular, (b) $\Delta^{\phi}(\mathscr{Q}_P) > 0 \Leftrightarrow H(P, \phi) > 0$; (c) $\Delta^{\phi}(\mathscr{Q}) = 0$ for all $\mathscr{Q} \subset \phi\mathscr{Q} \Leftrightarrow H = 0$; and (d) $\Delta^{\phi}(\mathscr{Q}_P) > 0$ for all nontrivial finite partitions $P \Leftrightarrow \phi$ is a K automorphism. Moreover, if ϕ is ergodic and $H(\phi) < \infty$, then (e) $\Delta^{\phi} = kf(H(\phi))$, where f is defined by

$$f(x) = \begin{cases} 0 & \text{if } \; x = 0 \\ \log(n+1) & \text{if } \; \log n < x \leqslant \log(n+1) \qquad (n = 1, 2, 3, \ldots) \end{cases}$$

and (f) $\Delta^{\phi}_{\mathrm{as}} = kH(\phi)$.

Proof. Parts (e) and (f) will be proven in a separate publication.[18] Part (a) follows from the observation that for $\mathscr{Q} \subset \phi\mathscr{Q}$

$$\Delta^{\phi}(\mathscr{Q}) > 0 \Leftrightarrow \mathscr{Q} \neq \phi^{-1}\mathscr{Q} \Leftrightarrow H(\mathscr{Q} \| \phi^{-1}\mathscr{Q}) \neq 0$$

Since $H(P, \phi) = H(\mathscr{Q}_P \| \phi^{-1}\mathscr{Q}_P)$, (b) follows from (a). Since $H(\phi) = \sup_{\mathscr{Q} \subset \phi\mathscr{Q}} H(\mathscr{Q} \| \phi^{-1}\mathscr{Q})$,[24] (c) also follows from (a). Part (d) follows from (b) using the fact that K-automorphisms ϕ have completely positive entropy; this means just that $H(P, \phi) > 0$ for all nontrivial P. ∎

[5] $H(\mathscr{Q} \| \phi^{-1}\mathscr{Q})$ is the conditional entropy of $\mathscr{Q}$ given $\phi^{-1}\mathscr{Q}$.[24]

Normally, in statistical mechanics we would expect to find $H \gg 1$, in which case the theorem says that if $\mathscr{C}$ is suitably chosen, then $\Delta^{\phi}(\nu, \mathscr{C})$ can be as large as H but very little larger.

Another question which can be partly answered using the Kolmogorov–Sinai entropy invariant is whether a σ-algebra $\mathscr{C}$ of the form $\mathscr{C}_P$ for some finite partition P has either of the properties considered in Corollary 2.2:

i. $\phi_s \mathscr{C} \uparrow \hat{1}$ as $s \to \infty$
ii. $\phi_s \mathscr{C} \downarrow \hat{0}$ as $s \to -\infty$

Theorem 4. If ϕ_1 is ergodic, then (i) $H < \infty$ if and only if there exists a finite partition P for which

$$\phi_s \mathscr{C}_P \uparrow \hat{1} \qquad \text{as } s \to \infty \qquad (s \in \mathbb{Z})$$

(ii) $H > 0$ if and only if there exists a nontrivial finite partition P for which

$$\phi_s \mathscr{C}_P \downarrow \hat{0} \qquad \text{as } s \to -\infty \qquad (s \in \mathbb{Z})$$

Proof. (i) By Krieger's theorem (Ref. 11, Sec. 9.7; Ref. 19) the condition $H < \infty$ implies that the dynamical system has a finite generator, which can be used as P. And if a finite P exists, then the definition of KS entropy implies

$$H < -\sum_{A \in P} \mu(A) \log \mu(A) < \infty$$

(ii) By Sinai's weak isomorphism theorem (Ref. 20; Ref. 11, Chap. 8; Ref. 9, p. 92) every ergodic automorphism with positive H has a Bernoulli factor of full KS entropy. Since every Bernoulli system is a K-system (Ref. 15, p. 32; Ref. 11, Sec. 7.15), any finite partition contained in this factor will have trivial tail and can therefore be used as P. On the other hand, if $H = 0$, then for all nontrivial P the algebra $\phi_s \mathscr{C}_P$ is independent of s and therefore cannot tend to $\hat{0}$. ■

For any classical system Kushnirenko's theorem (Refs. 21; 15, p. 46) tells us that $H < \infty$, and so if the system is ergodic, a finite P satisfying condition (i) can be found; but to show that $H > 0$ is more difficult.

9. DISCUSSION

The main result of this paper is a method of defining a time-dependent entropy for a nonequilibrium measure (i.e., a nonequilibrium ensemble) in a general dynamical system. This definition has the main properties one would want of nonequilibrium entropy. It does not decrease with time (proved in Corollary 2.2); under suitable conditions [specified in part (ii) of

Corollary 2.2] it tends to the thermodynamic equilibrium entropy as time proceeds; and the entropy of a composite system consisting of two independent parts is easily shown (although we have not given the details in this paper) to be the sum of their individual entropies.

Our definition of entropy employs a sub-σ-algebra $\mathcal{Q}$ of measurable sets satisfying the condition (2.1), $\phi_t\mathcal{Q} \supset \mathcal{Q}$, $t > 0$. According to the interpretation of the algebra $\mathcal{Q}$ proposed in Section 2, its elements correspond to the various possible propositions about the future observable behavior of the system. The condition $\phi_t\mathcal{Q} \supset \mathcal{Q}$ represents the fact that as time proceeds the set of observational opportunities that are still in the future decreases, and our entropy nondecrease law (Corollary 1.2) corresponds to the way this loss of observational opportunities progressively reduces the amount of detailed information that can still be obtained about any specific ensemble by means of observations made in the future.

Two further conditions on $\mathcal{Q}$ are also considered in the paper. One, condition (i) of Corollary 2.2, has the physical interpretation that a sequence of observations stretching infinitely far both into past and future will determine the microscopic state as accurately as we please; it has the consequence (Corollary 2.2) that as $t \to -\infty$ our entropy tends to the fine-grained entropy. The other condition, (ii) of Corollary 2.2, has the physical interpretation that a sequence of observations starting infinitely far in the future will give no significant information about the microscopic state; this condition has the consequence that as $t \to \infty$ our entropy approaches the equilibrium entropy. In a K-system, both conditions are satisfied (if $\mathcal{Q}$ is chosen properly), and many of the mechanical systems whose ergodic properties have been studied in detail, such as the various types of "billiards" studied by Sinai, Bunimovich,[22] and others, are K-systems. If condition (i) is violated, but (ii) holds, then the system has a factor which is a K-system, constructed by treating the elements of the measurable partition associated with $\bigvee_{-\infty}^{\infty} \phi_t\mathcal{Q}$ (the "physical" algebra, of events observable in the past, present, or future) as the points in a new phase space, and so the main results are still essentially the same. On the other hand, Theorem 4 implies that if ϕ_1 is ergodic, then the only way that condition (ii) can fail for all $\mathcal{Q} \subset \phi_1\mathcal{Q}$ is by the vanishing of the Kolmogorov–Sinai entropy invariant.

The Kolmogorov – Sinai entropy is also closely related to the rate at which our entropy can increase; a more detailed treatment of this question, including the proof of Theorem 3, parts (e) and (f), will be given in a separate paper.

We are interested particularly in σ-algebras $\mathcal{Q}$ of the form $\mathcal{Q}_P$ for some finite P, which may be interpreted in terms of observational equivalence. Theorem 4 says that σ-algebras of this form which satisfy (i) or (ii) just mentioned exist in any (discrete-time) ergodic dynamical system whose Kolmogorov–Sinai entropy invariant is finite and positive. There is no need to assume, as is done, for example, by Penrose (Ref. 14, Sec. 5), that the observational states form a Markov chain, although the result given in Section 6 shows that our results reduce to the usual ones in the special case where the states do form a Markov chain.

There is an apparent contradiction—the "paradox of irreversibility"—between the time-reversal asymmetry of entropy increase and the time-reversal symmetry of Newtonian dynamical systems. Of course, for our entropy the contradiction disappears when we realize that the algebra $\mathscr{R}$ used in the definition of entropy is not symmetric under time reversal. The suggestion that the paradox of irreversibility can be avoided by using a nonsymmetric definition of entropy was made by Prigogine, George, Henin, and Rosenfeld,[23] but they followed it up in a different way; they eschew coarse graining and instead seek to express the entropy in terms of the expectation of a suitable dynamical variable. An interesting recent application of this idea is the work of Misra,[4] who constructs a dynamical variable which can be interpreted as the time and is nondecreasing for that reason.

ACKNOWLEDGMENT

We are grateful to J. L. Lebowitz for a discussion about the baker's transformation example in Section 6, which prompted this work.

REFERENCES

1. A Wehrl, *Rev. Mod. Phys.* **50**:221–260 (1978).
2. O. Penrose, *Rep. Prog. Phys.* **42**:1937–1997 (1979).
3. I. Prigogine, *Nature* **246**:67–71 (1973).
4. B. Misra, *Proc. Nat. Acad. Sci. USA* **75**:1627–1631 (1978).
5. R. C. Tolman, *The Principles of Statistical Mechanics* (Oxford University Press, Oxford, 1938), Sec. 51.
6. S. T. Ali and G. G. Emch, *J. Math. Phys.* **15**:176–182 (1974).
7. L. M. Morato, *Int. J. Theor. Phys.* **16**:707–713 (1977).
8. E. Prugovečki, *J. Math. Phys.* **17**:517–523 (1976).
9. P. Billingsley, *Ergodic Theory and Information* (Wiley, New York, 1965), p. 112.
10. J. L. Doob, *Stochastic Processes* (Wiley, New York, 1953).
11. M. Smorodinsky, *Ergodic Theory, Entropy*, Springer Lecture Notes in Mathematics No. 214 (Springer, New York, 1971), Sec. 3-6.
12. Ya. G. Sinai, in *Statistical Mechanics, Foundations and Applications* (Proc. IUPAP Meeting, Copenhagen, 1966), T. Bak, ed., pp. 559–573.
13. P. R. Halmos, *Lectures on Ergodic Theory* (Mathematical Society of Japan, 1965), p. 9.
14. O. Penrose, *Foundations of Statistical Mechanics* (Pergamon, New York, 1970), p. 122.
15. V. I. Arnold and A. Avez, *Ergodic Problems of Classical Mechanics* (Benjamin, New York, 1968).
16. W. Pauli, Uber das H. Theorem von Anwachsen der Entropie von Standpunkt der neuen Quantenmechanik, in *Probleme der Moderne Physik*, P. Debye, ed. (Hirzel, Leipzig, 1927).

17. G. Lindblad, *Commun. Math. Phys.* **40**:147–151 (1975).
18. S. Goldstein, Entropy increase in dynamical systems, *Israel J. Math.* (to appear).
19. W. Krieger, *Trans. Am. Math. Soc.* **199**:453–464 (1970).
20. Ya. G. Sinai, *Dokl. Akad. Nauk USSR* **147**:797–800 (1962).
21. A. G. Kushnirenko, *Dokl. Akad. Nauk USSR* **161**(1):37–38 (1965).
22. Ya. G. Sinai, *Dokl. Akad. Nauk USSR* **153**:1261–1264 (1963); L. A. Bunimovich, *Commun. Math. Phys.* **65**:295–312 (1979).
23. I. Prigogine, C. George, F. Henin, and L. Rosenfeld, *Chemica Scripta* **4**:5–32 (1973).
24. W. Parry, *Entropy and Generators in Ergodic Theory* (Benjamin, New York, 1969), Chap. 4, Sec. 1.