Journal of Statistical Physics, Vol. 77, Nos. 1/2, 1994

Canonical Nonequilibrium Ensembles and Subdynamics

Peter V. Coveney[1] and Allan K. Evans[2]

Received October 28, 1993; final January 10, 1994

A method for constructing a canonical nonequilibrium ensemble for systems in which correlations decay exponentially has recently been proposed by Coveney and Penrose. In this paper, we show that the method is equivalent to the subdynamics formalism, developed by Prigogine and others, when the dimension of the subdynamic kinetic subspace is finite. The comparison between the two approaches helps to clarify the nature of the various operators used in the Brussels formalism. We discuss further the relationship between these two approaches, with particular reference to a simple discrete-time dynamical system, based on the baker's transformation, which we call the baker's urn.

KEY WORDS: Nonequilibrium statistical mechanics; irreversibility; subdynamics; exponential decay of correlations.

1. INTRODUCTION

A central problem in nonequilibrium statistical mechanics is the derivation of irreversible macroscopic behavior from reversible microscopic dynamics. For isolated systems, the microscopic equations of motion are time-reversal symmetric; the fact that macroscopic systems are time asymmetric indicates that the ensembles or probability distributions which we use to describe them are themselves not symmetric under time reversal. Thus, as Oliver Penrose stated in his 1979 review on the foundations of statistical mechanics, "it would be a considerable advance in nonequilibrium statistical mechanics if someone succeeded in characterising the set of physically possible nonequilibrium ensembles."[1]

[1] Schlumberger Cambridge Research, High Cross, Cambridge CB3 OHG, United Kingdom.
[2] Cavendish Laboratory, University of Cambridge, Cambridge CB3 OHE, United Kingdom.

0022-4715/94/1000-0229$07.00/0

Related to this problem is the fact that it is rare to work with the full N-particle distribution function for a many-body system. Instead, some "reduced" description, such as the one-particle distribution function in Boltzmann's equation, is used, and an irreversible kinetic equation is derived for this simpler description. The validity of methods for describing the time evolution of reduced ensembles, and the accompanying interpretations, have been the subject of controversy since the time of Boltzmann.[1,2]

A group based in Brussels and led by Prigogine has invested several decades of work in developing a general method for selecting a representative, reduced ensemble and following its evolution as time passes.[3–6] This method is known as the subdynamics, or "Brussels" formalism. One would expect that the class of time-dependent ensemble densities selected in this manner should be related to the class of physically possible ensembles.[1] The Brussels method has been used to derive kinetic equations[4] and also to treat the quantum mechanical decay problem[7] and discrete-time dynamical systems.[8,9] However, several quite different formulations of the theory have appeared over the years, and the complexity of the formalism has made it difficult to establish rigorous results.

Recently, Penrose and Coveney[10] have proposed a rigorous measure-theoretic method of constructing a time-asymmetric "canonical nonequilibrium ensemble" (CNEE). The use of the word "canonical" here is intended to convey the idea that the ensemble furnishes a simple, standard, reduced description which represents the most important features of the approach to equilibrium. These authors illustrated their ideas using a model discrete-time system, which they named the "pastry-cook's transformation."

In a subsequent paper, Evans and Coveney[11] showed that for the pastry-cook model, the subdynamics formalism could be used to construct a representative ensemble identical to the CNEE. They also located the dynamical resonances, discussed by Ruelle[12] and others,[13,14] for the pastry-cook model, showing that the rate of decay of correlation functions at long times given by the resonances was the same as that given by the subdynamics method.

In this paper, we extend these results. After defining in Section 2 a projection operator which maps densities onto the CNEE, in Section 3 we briefly summarize one version of the Brussels formalism. In Section 4, we show that under certain conditions, which are no more than are necessary to guarantee the validity of the subdynamics method, a finite-dimensional kinetic subspace in subdynamics is equivalent to the CNEE of Penrose and Coveney.

2. THE CANONICAL NONEQUILIBRIUM ENSEMBLE

We shall work with a classical discrete-time dynamical system, which is defined by a phase space Γ and a mapping T from Γ onto itself. Given a starting position x_0 in Γ, the state of the system at time t is $x_t = T^t x_0$. To make progress in statistical mechanics we must describe the time-evolution of probability distributions $\rho(x)$. In principle, the distribution might be any integrable function. However, as we will see later, for some purposes we must impose restrictions on ρ.

As the map T moves states forward in time, so the Frobenius–Perron operator U moves distributions forward in time.[15] U is defined by the equation

$$U\rho(x) = \int_\Gamma \delta(x - Ty)\, \rho(y)\, dy \tag{1}$$

where $\delta(x)$ is the Dirac delta function. If ρ is the distribution at time t, then $U\rho$ is the distribution at time $t+1$. The long-time behavior in statistical mechanics is determined by the behavior of $U^t\rho$ for large t. We will assume that the map T is ergodic, so that averages over long times converge and define a natural equilibrium measure π on the phase space.[16]

We now introduce an asymptotic operator $\tilde{\Pi}$,[11] in the spirit of the Brussels projection operator Π. The Brussels Π extracts from a probability distribution ρ the component $\Pi\rho$, which is associated with a particular independent mode in the subdynamics of the system. The new operator $\tilde{\Pi}$ extracts the component $\tilde{\Pi}\rho$, which decays exponentially with the decay rate λ ($|\lambda| < 1$) of the slowest mode in the system. This is the more important part of the canonical nonequilibrium ensemble (CNEE) of Coveney and Penrose. The other component of their CNEE is simply a uniform distribution which is unchanged as time passes; we omit this component here. We define the asymptotic operator as follows: for $x \in \Gamma$,

$$\tilde{\Pi}\rho(x) := \text{weak} \lim_{t \to \infty} \lambda^{-t}\{U^t\rho(x) - P^\pi\rho(x)\} \tag{2}$$

where P^π is a projection operator which maps onto the space spanned by the equilibrium distribution $\pi(x)$. Note that this definition differs slightly from that given by Coveney and Evans[11] in that here we omit the second term $P^\pi\rho(x)$. This will allow us to express the results of Section 4 in a simple form.

For systems which are mixing, any matrix element of the operator

$U^r - P^\pi$ tends to zero: that is, for mixing systems and measurable functions ϕ, ρ,

$$\lim_{n \to \infty} \langle \phi, [U^n - P^\pi] \rho \rangle = 0 \tag{3}$$

where the brackets denote the usual scalar product of the functions. However, for $\tilde{\Pi}$ to be well-defined, we need a stronger condition: that the rate of convergence to zero in (3) is at least as fast as λ^n.

The "weak limit" in (2) means that we confine ourselves to talking about matrix elements of the operator, rather than the operator itself. The operator $\tilde{\Pi}$ is therefore well-defined only when it acts on a distribution ρ such that matrix elements of the form

$$\langle \phi, \tilde{\Pi}\rho \rangle = \int_\Gamma \phi(x)\, \tilde{\Pi}\rho(x)\, dx \tag{4}$$

converge. This means that the function $\phi\tilde{\Pi}\rho$ must be integrable. The conditions which this imposes on ϕ and ρ will depend upon the details of the dynamical system.

The definition of the asymptotic operator [Eq. (2)] is simply motivated. We assume that the dynamics of probability distributions can be separated into a number of exponentially decaying modes, together with a component parallel to the equilibrium distribution, which is time-invariant. After a large time t, for almost any starting distribution ρ, modes other than the slowest will make a negligible contribution to $U^t\rho$. We subtract the component of ρ parallel to the equilibrium distribution π. Then we extrapolate back to time zero, using the decay constant λ of the slowest mode; this procedure supplies the amplitude of the most slowly decaying mode at time zero.

The crucial assumption here is that the dynamics of probability distributions can be separated into independent, exponentially decaying components, as in subdynamics.[17] Recent work by Bandtlow and Coveney[18] has shed light on the subject, although their condition for the separation of long-lived modes is not easy to apply in practice, as it depends on the topology used for the space of probability distributions.

From the definition (2), we can immediately see that $\tilde{\Pi}$ is an idempotent projection operator ($\tilde{\Pi}^2 = \tilde{\Pi}$), and that it commutes with the evolution operator U. These two features capture the essence of subdynamics as defined by the Brussels group.[4] $\tilde{\Pi}$ projects onto a subspace of probability distributions which is invariant under the time evolution. Since this is exactly the purpose of the projection operators defined by the Brussels school, it is natural to ask whether the operators are in fact the same. In the next two sections we consider this question in more detail.

3. THE SUBDYNAMICS FORMALISM

This section discusses one version of the subdynamics formalism as it applies to discrete-time dynamical systems. It is by no means a complete treatment, and it is only one of several formulations of the theory.

The resolvent-based formalism we present here is based on the continuous-time version in the book by Balescu.[4] For a concise summary, see Coveney and Penrose.[17] A "time-dependent" formulation not involving the resolvent has been developed by Coveney and others[19] and more recently the theory has been based on a set of commutation or "intertwining" relations.[6,9] All these versions are equivalent.

More details of the continuous-time formalism and applications in statistical mechanics can be found in the book by Balescu.[4] For the discrete-time case, see recent papers by Bandtlow and Coveney[18] and Hasegawa and Saphir.[8]

The Brussels approach is based on the idea of subdynamics. A time-dependent ensemble density $\rho(t)$ (a phase space density in classical mechanics or a density matrix in quantum mechanics) is separated into several independently evolving components by a complete set of projection operators Π^i. Each projector defines a subspace of densities which is invariant under the evolution of the system with time. More precisely, the projection operators each commute with the evolution operator.

In statistical mechanics, the kinetic component $\Pi\rho$ is the most important, as it is associated with the long-time evolution of the system under consideration. Kinetic equations such as Boltzmann's are derived as approximations to an exact kinetic equation which is obeyed by part of this kinetic component. The other invariant subspaces contain information in which we are normally not interested, such as correlations between several particles. For systems in thermodynamic equilibrium and for nonequilibrium stationary states, the entire description is supposed to lie within the kinetic subspace which Π defines.[4] In the current paper, the single Π subspace which we construct is intended to be analogous to the kinetic subspace in statistical mechanics.

Recently, attention has been directed toward abstract dynamical systems, such as the baker transformation.[8,9] The study of abstract dynamical systems has clarified some fundamental features of the subdynamics formalism. The choice of the (vector) space of ensemble densities is now recognized to be important. The operators occurring in the formalism are taken to act in a "rigged" Hilbert space which is not self-dual, rather than in a true Hilbert space; it is maintained that this provides the time asymmetry needed for a description of irreversibility.[9]

Moreover, it has become clear that the modes studied by the Brussels

group[8] are closely related to the resonances uncovered by Ruelle and others[12,13] which are poles in the Fourier transforms of correlation functions. These poles give the decay rates of exponentially decaying modes in subdynamics.

The starting point for the "resolvent-based" formulation of subdynamics is the resolvent operator, defined by

$$R(z) = \frac{1}{z - U} = \frac{1}{z} \sum_{t=0}^{\infty} U^t z^{-t} \tag{5}$$

$R(z)$ is the Laplace transform of U; by inverting the transform, we can reconstruct the dynamics of the system.

The aim is to construct an operator $\Sigma(z)$ which picks out the longest-lived mode in the system's evolution. The hope is that $\Sigma(z)$ will be a simpler object than $R(z)$, but will still give the correct dynamical behavior as $t \to \infty$. This is a familiar strategy in statistical mechanics; for example, in the theory of Brownian motion, by isolating the "slow variables" associated with the motion of a massive test particle and averaging over the other degrees of freedom of the system, one extracts a description of the motion which is valid for large time scales.

We define an orthogonal projection operator P which projects from the full description of the ensemble given by $\rho(x)$ to a reduced description which, though incomplete, determines the values of our "slow" variables. For the case of Brownian motion, P might be chosen so that if $\rho(x)$ is an N-particle distribution function for the Brownian particle and its surrounding gas molecules, then $P\rho(x)$ is a one-particle distribution function for the Brownian particle only. In the case of quantum mechanics, P is usually defined to pick out the diagonal part of the density matrix.

The Brussels decomposition of $R(z)$ enables us to construct (with some effort) the full resolvent $R(z)$ from its projection $PR(z)P$. In fact, it can be shown that[18]

$$R(z) = [P + C(z)]\, PR(z)\, P[D(z) + P] + S(z) \tag{6}$$

The operators $C(z)$ and $D(z)$ are known as creation and destruction operators.[4] They are defined as follows:

$$C(z) = \frac{1}{z - QUQ} QUP = \sum_{r=1}^{\infty} \left(\frac{QU}{z}\right)^r P \tag{7}$$

and

$$D(z) = PUQ \frac{1}{z - QUQ} = P \sum_{r=1}^{\infty} \left(\frac{UQ}{z}\right)^r \tag{8}$$

where $Q = 1 - P$ projects onto the subspace orthogonal to the P subspace. The operator $S(z)$ is known as the reduced resolvent, and is defined by

$$S(z) = Q \frac{1}{z - QUQ} Q = \sum_{r=1}^{\infty} Q(UQ)^{r-1} z^{-r} \tag{9}$$

In the rigorous justification of this version of the subdynamics formalism[18] an important point is that if the projector P is chosen so that all the "slow variables" are determined by the P subspace projection of the ensemble density, then matrix elements of $(QUQ)^n$ decay exponentially with n at a rate strictly faster than the slowest mode in the system. This means that the operators $C(z)$, $D(z)$, and $S(z)$, which share the denominator $z - QUQ$, are each regular for $|z| \geqslant |\lambda|$, where λ is the decay rate of the slowest mode.[11,18]

Once a decaying mode has been found by locating a pole of the simpler resolvent $PR(z)\,P$, Eq. (6) can be used to write down an expression for the decaying component of the full resolvent corresponding to that pole. This is useful because modes constructed in this way are truly independent, evolving with time separately from the rest of the dynamics. This is the idea behind the term "subdynamics."[4]

Given an asymptotic form $P\Sigma(z)\,P$ for the partial resolvent $PR(z)\,P$, from Eq. (6) we obtain an asymptotic form for the full resolvent:

$$\Sigma(z) = [P + C(\lambda)]\, P\Sigma(z)\, P[D(\lambda) + P] \tag{10}$$

where λ is the location of the pole in the partial resolvent. $C(z)$ and $D(z)$ are replaced by $C(\lambda)$ and $D(\lambda)$ because it is the residue of the pole at $z = \lambda$ which determines the behavior of the component, and C and D are regular at this point. The reduced resolvent $S(z)$ plays no part in the long-time behavior of the system, because of the rapid decay of $(QUQ)^n$.

$\Sigma(z)$ is the resolvent operator for a single independent mode in the subdynamics. From it we can obtain a projection operator Π, which, acting on any distribution ρ, gives us the component $\Pi\rho$ associated with the mode. Since $\Pi R(z) = R(z)\,\Pi = \Sigma(z)$, from the definition (5) of the resolvent we have

$$\Pi = \lim_{z \to \infty} z\Sigma(z) \tag{11}$$

Therefore

$$\Pi = [P + C(\lambda)]\, P\Pi P[D(\lambda) + P] \tag{12}$$

After we have constructed an operator $\Sigma(z)$ for the longest-lived mode in the system, the time-dependent evolution operator for this mode can be obtained from the inverse transform:

$$\Sigma_t = \frac{1}{2\pi i} \oint z^t \Sigma(z)\, dz \tag{13}$$

where the contour encloses the unit circle. The long-time behavior of a correlation function

$$g_{\phi\psi} = \langle \phi, U^t \psi \rangle - \langle \phi, 1 \rangle \langle 1, \psi \rangle \tag{14}$$

is then given by

$$\lim_{t \to \infty} g_{\phi\psi}(t) = \langle \phi, \Sigma_t \psi \rangle \tag{15}$$

By using a complete set of projection operators P_i, each projecting onto a finite-dimensional subspace, and extending the asymptotic form of each partial resolvent in this way, Hasegawa and Saphir[8] were able to give a complete spectral decomposition of the baker transformation, including eigenvalues and eigenstates for the Frobenius–Perron operator. An important conclusion of their work was that a full spectral description of the dynamics of probability distributions cannot be achieved in the Hilbert space $\mathscr{H}$ of square-integrable functions over the phase space. Instead, the probability distributions are restricted to a subset of $\mathscr{H}$, while the dual space is enlarged to include a larger class of functionals than the dual of $\mathscr{H}$. The space of probability distributions is then spanned by a countable set of exponentially decaying eigenstates of U. This "rigged" Hilbert space is similar to those used by Böhm[20] to treat exponential decay in quantum mechanics.

Another operator which is important in the theory of subdynamics is the collision operator $\psi(t)$, which can be thought of as a measure of the influence of the Q subspace on the dynamics in the P subspace:

$$\psi(t) = PU(QU)^t P \tag{16}$$

Through its Laplace transform,

$$\tilde{\psi}(z) = \sum_{t=0}^{\infty} PU(QU)^t P z^{-t} \tag{17}$$

the collision operator is related to the P subspace resolvent:

$$PR(z)\, P = P \frac{1}{Pz - PUP - \tilde{\psi}(z)} P \tag{18}$$

Coveney and Penrose[17] recently gave a set of sufficient conditions which, when applied to the collision operator, guarantee the validity of part of the continuous-time subdynamics formalism. They proved that if $\psi(t)$ is bounded above in norm by an exponentially decaying function of t, and the P subspace is finite-dimensional, then the Brussels formalism allows the separation of an exponentially decaying mode which gives the long-time behavior in the P subspace.

4. THE RELATIONSHIP BETWEEN SUBDYNAMICS AND THE CANONICAL NONEQUILIBRIUM ENSEMBLE

In this section, we show that when the P subspace is finite-dimensional, we can construct from the asymptotic operator of Section 2 a set of operators identical with the Brussels operators of the last section.

First, we define operators $\tilde{C}$ and $\tilde{D}$ analogous to $C(\lambda)$ and $D(\lambda)$ by solving the equations

$$[P+\tilde{C}]\,P\tilde{\Pi}P=\tilde{\Pi}P \tag{19}$$

and

$$P\tilde{\Pi}P[\tilde{D}+P]=P\tilde{\Pi} \tag{20}$$

to give

$$\tilde{C}=\tilde{\Pi}P[P\tilde{\Pi}P]^{-1}-P \tag{21}$$

and

$$\tilde{D}=[P\tilde{\Pi}P]^{-1}P\tilde{\Pi}-P \tag{22}$$

Then we define

$$\begin{aligned}\bar{\Pi}&=[P+\tilde{C}]\,P\tilde{\Pi}P[\tilde{D}+P]\\&=\tilde{\Pi}P[P\tilde{\Pi}P]^{-1}P\tilde{\Pi}\end{aligned} \tag{23}$$

This operator is identical to the Brussels Π, as we will now show.

We use the identities

$$P\Pi P=P\,\frac{1}{P+D(\lambda)\,C(\lambda)}\,P \tag{24}$$

(see, for example, Antoniou and Tasaki[9]),

$$P + C(z) = R(z)\, P \frac{1}{PR(z)\, P} \tag{25}$$

and

$$D(z) + P = \frac{1}{PRP(z)} PR(z) \tag{26}$$

The last two follow from Eq. (6). For example, Eq. (25) can be obtained by multiplying (6) on the right by $P/PR(z)\, P$. Equations (25) and (26) allow us to write Eq. (24) in the form

$$\Pi = \text{weak} \lim_{z \to \lambda} \left\{ R(z)\, P \frac{1}{PR(z)\, P} P\Pi P \frac{1}{PR(z)\, P} PR(z) \right\} \tag{27}$$

Taking the limit $z \to \lambda$ is necessary because both $PR(z)\, P$ and $R(z)$ have a pole at $z = \lambda$, so neither is well-defined at that point. Since both poles are of the same order, the quotient tends to a well-defined limit, as we now show by expanding the resolvents in Eq. (27) using the definition (5). For simplicity, we assume that the P subspace is one-dimensional. A matrix element of the operator $R(\lambda)/PR(\lambda)\, P$ can be written as

$$\left\langle \phi, \frac{R(\lambda)\, P}{PR(\lambda)\, P} c \right\rangle = \lim_{z \to \lambda} \frac{\sum_{r=0}^{\infty} U^r_{\phi c} z^{-r}}{\sum_{r=0}^{\infty} U^r_{cc} z^{-r}} \tag{28}$$

where c is a function spanning the one-dimensional P subspace and U^r_{ab} denotes the matrix element $\langle a, U^r b \rangle$.

We now separate each matrix element into an asymptotic part, which decays with the slowest mode's decay constant λ, and a nonasymptotic part, which decays more rapidly:

$$U^r_{ab} = \lambda^r \tilde{\Pi}_{ab} + \Delta^{(r)}_{ab} \tag{29}$$

Provided the subdynamics decomposition is valid, the nonasymptotic part Δ decays as $r \to \infty$ at a rate faster than the slowest mode. When we evaluate the infinite sums, we find

$$R_{ab}(z) = \frac{1}{z} \sum_{r=0}^{\infty} U^r_{ab} = \frac{\tilde{\Pi}_{ab}}{z - \lambda} + \frac{1}{z} \sum_{r=0}^{\infty} \Delta^{(r)}_{ab} z^{-r} \tag{30}$$

As $z \to \lambda$, the second term on the right-hand side of this equation becomes negligible compared to the first term and we get

$$\begin{aligned}\langle \phi, (P + C(\lambda))\, c\rangle &= \lim_{z\to\lambda} \left\langle \phi, \frac{R(\lambda)}{PR(\lambda)\,P}\, c \right\rangle \\ &= \frac{\tilde{\Pi}_{\phi c}}{\tilde{\Pi}_{cc}} \\ &= \langle \phi, (P + \tilde{C})\, c\rangle \end{aligned} \tag{31}$$

This shows that $\tilde{C} = C(\lambda)$. A similar calculation starting with Eq. (26) shows that $\tilde{D} = D(\lambda)$, so all that remains to establish the identity of Π and $\bar{\Pi}$ is to show that $P\tilde{\Pi}P = P\Pi P$.

Using (24) and the results of the last paragraph, we can write

$$P\Pi P = P \frac{1}{P + \tilde{D}\tilde{C}} P \tag{32}$$

But

$$\begin{aligned}\tilde{D}\tilde{C} &= ([P\tilde{\Pi}P]^{-1} P\tilde{\Pi} - P)(\tilde{\Pi}P[P\tilde{\Pi}P]^{-1} - P) \\ &= [P\tilde{\Pi}P]^{-1} P\tilde{\Pi}^2 P[P\tilde{\Pi}P]^{-1} - P\tilde{\Pi}P[P\tilde{\Pi}P]^{-1} \\ &\quad - [P\tilde{\Pi}P]^{-1} P\tilde{\Pi}P + P \\ &= [P\tilde{\Pi}P]^{-1} - P \end{aligned} \tag{33}$$

Therefore

$$P \frac{1}{P + \tilde{D}\tilde{C}} P = P\tilde{\Pi}P \tag{34}$$

Hence $P\tilde{\Pi}P = P\Pi P$, and we have established the identity of the two expressions for the projector: $\bar{\Pi} = \Pi$. In other words,

$$\Pi = \tilde{\Pi}P[P\tilde{\Pi}P]^{-1} P\tilde{\Pi} \tag{35}$$

For the proof outlined above to be meaningful, $[P\tilde{\Pi}P]$ must be invertible. Since both P and $\tilde{\Pi}$ are projection operators, $[P\tilde{\Pi}P]^{-1}$ is well-defined provided that whenever $\tilde{\Pi}\rho = 0$, $P\rho = 0$ also. In physical terms, all the variables which the reduced, P subspace description of the system determines must be coupled to the slowest-decaying mode.

The extension of the above to a P subspace of finite dimension greater than one is straightforward. The only important difference is that $\tilde{\Pi}_{cc}$ is

replaced by an operator with finite-dimensional range, which can be expressed as a matrix with respect to some set of functions which spans the subspace. The case where the P subspace is infinite-dimensional is more difficult to treat; we expect that the results of this section will not generalize to this case, because this would make a treatment of nonexponential decay in the subdynamics formalism impossible.

Equation (35) expresses an important relationship between the two operators Π and $\tilde{\Pi}$, but it does not show that they are identical. One case where the two are identical is when the dynamical system satisfies the condition known as "Axiom A" (for a definition, see Ruelle[21]), and both P and $\tilde{\Pi}$ are one-dimensional. In the case of Axiom A systems,[13] the long-time behavior of correlation functions is governed by an equation of the form

$$\lim_{t \to \infty} g_{\phi\psi}(t) = \langle \phi, A \rangle \langle B, \psi \rangle \alpha^t \tag{36}$$

where α is a constant with $|\alpha| < 1$, and A and B are distributions. Hence

$$\langle \phi, \tilde{\Pi}\psi \rangle = \langle \phi, A \rangle \langle B, \psi \rangle \tag{37}$$

This, together with Eq. (35), yields the following expression for matrix elements of Π:

$$\begin{aligned} \langle \phi, \Pi\psi \rangle &= \langle \phi, A \rangle \langle B, P[P\tilde{\Pi}P]^{-1} PA \rangle \langle B, \psi \rangle \\ &= \langle \phi, A \rangle \langle B, \psi \rangle \langle B, P[P\tilde{\Pi}P]^{-1} PA \rangle \\ &= \langle \phi, A \rangle \langle B, \psi \rangle \cdot \text{const} \\ &= \langle \phi, \tilde{\Pi}\psi \rangle \cdot \text{const} \end{aligned} \tag{38}$$

Since we have already established that $P\Pi P = P\tilde{\Pi}P$, this constant can be nothing but unity. Hence

$$\langle \phi, \Pi\psi \rangle = \langle \phi, \tilde{\Pi}\psi \rangle \tag{39}$$

for Axiom A systems where P and $\tilde{\Pi}$ are one-dimensional and $[P\tilde{\Pi}P]^{-1}$ exists.

5. THE BAKER'S URN

The dynamical system analyzed here is a simple extension of the well-known baker transformation, and shares several features in common with the pastry-cook transformation introduced in ref. 10. The baker transformation itself has been widely used because of its simplicity, and because it

shares several characteristics with the Hamiltonian systems which are the subject of statistical mechanics: it is a measure-preserving, ergodic K-automorphism and its dynamics is reversible.[16]

The phase space Γ_0 of the baker transformation is the unit square $[0, 1) \times [0, 1)$ in the (x, y) plane. The transformation is a bijective mapping T from Γ_0 onto itself, described by

$$T(x, y) = \begin{cases} \left(2x, \frac{y}{2}\right) & \text{if} \quad x < 1/2 \\ \left(2x - 1, \frac{y}{2} + \frac{1}{2}\right) & \text{if} \quad x \geqslant 1/2 \end{cases} \tag{40}$$

To visualize the transformation, imagine squashing the unit square to half its original height, also stretching it out horizontally so that its area is unchanged. Then break off the right-hand half of the rectangle and place it on top of the left half, to reform the square (see Fig. 1).

The baker transformation is easy to analyze because of its simple symbolic dynamics. Any point in the phase space can be described by a biinfinite string of binary digits $\{\omega_i; i \in \mathbf{Z}^+\}$. The infinite string $\omega_1 \omega_2 \omega_3 \cdots$

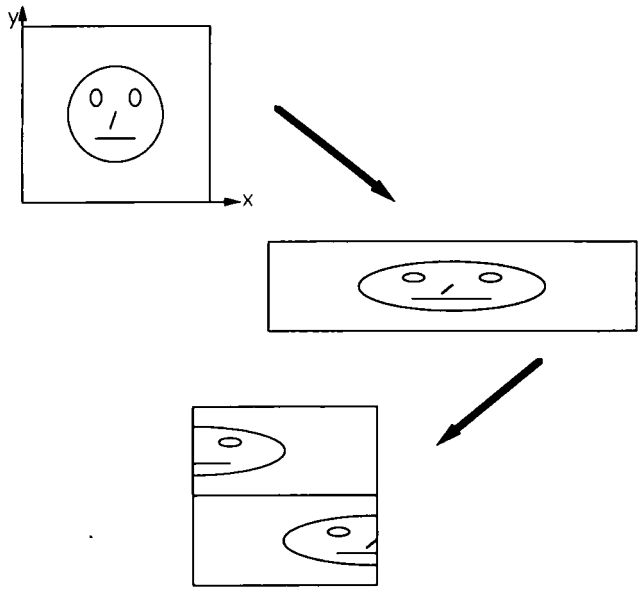

Fig. 1. The baker transformation.

is the binary expansion of x, and $\omega_0\omega_{-1}\omega_{-2}\cdots$ is y in binary. The baker transformation is then a shift of the string one place to the left:

$$T\omega_i = \omega_{i+1} \tag{41}$$

Once the transformation has been expressed in this form, it is possible to examine its ergodic properties in a straightforward way.[16]

We extend the baker transformation by introducing a new variable c, the "color," which can take one of two values: $c = 1$ (white) or $c = -1$ (black). We also define a small patch on the unit square: for example, χ could be the region $x < 1/4$; see Fig. 2. After each application of the ordinary baker transformation (40), we look to see whether the (x, y) coordinates of the system lie in the patch. If they do, then the color variable is changed. If not, then the color is left unchanged. This model is similar to, but distinct from, the pastry-cook transformation introduced by Penrose and Coveney.[10]

In the symbolic dynamics, a point in the phase space Γ is now described by the pair $x = (\omega, c)$, where ω is a biinfinite sequence of bits as before, and $c \in \{-1, 1\}$. The transformation is equivalent to the mapping

$$Tx = T(\omega, c) = (T\omega, (-1)^{\chi(T\omega)} c) \tag{42}$$

where $\chi(\omega)$ is a characteristic function defined to be 1 when ω is in χ and zero otherwise. We will take $\chi(\omega)$ to be a product of a finite number of adjacent binary digits:

$$\chi(\omega) = \omega_0\omega_1\cdots\omega_{k-1} \tag{43}$$

In this case, the color-changing patch is the region $x > 1 - 2^{-k}$.

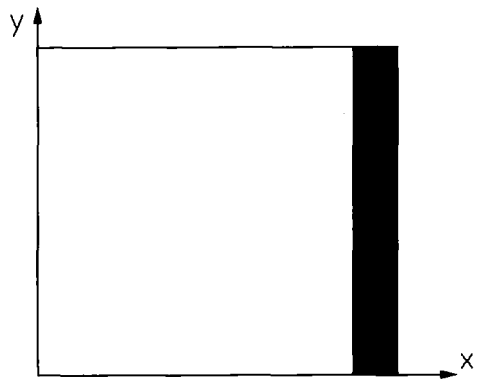

Fig. 2. Phase space for the baker's urn system. The color-changing patch is shaded black.

As one might expect, this dynamical system is mixing, and therefore ergodic. We will not give the proof of the mixing property here, since it is very similar to the proof for the pastry-cook's transformation given by Penrose and Coveney.[10]

Like the pastry-cook model and the Ehrenfests' urn,[22] this dynamical system can be thought of as a crude model of a gas contained in two vessels with convex walls and connected by a narrow constriction. One vessel is the $c = 1$ half of the phase space, the color-changing patch is the constriction, and the other vessel is the $c = -1$ region of the phase space. Equally, it might be taken to represent a single vessel containing two distinct molecular species which interconvert when the molecules hit the color-changing patch.

In the former case, if most of the gas were concentrated in the left-hand vessel, the system would slowly relax toward equilibrium as the concentrations equalized. A similar interpretation of the approach to equilibrium holds in the latter case. In our system, if an ensemble of points in the phase space starts with an uneven distribution of colors, the color distribution relaxes, through successive applications of the transformation, toward a fine-grained mixture of black and white—almost a uniform gray.

To make further progress with our analysis, we must describe the evolution of these ensembles of phase-space points. An ensemble is described by a probability distribution $\rho(x)$, and its evolution is given by the Frobenius–Perron[15] operator U generated by the transformation T. This has a particularly simple form for our measure-preserving, bijective map:

$$U\rho(x) = \rho(T^{-1}x) = \rho(T^{-1}\omega, (-1)^{\chi(\omega)} c) \tag{44}$$

6. *P*-SUBSPACE DESCRIPTION OF THE BAKER'S URN

We now specify a one-dimensional projection operator P, which projects the full dynamics onto the subspace spanned by the distribution c. Acting on an arbitrary function of the phase-space variables ω and c, P is defined as follows:

$$P\rho(\omega, c) = c \int_\Gamma \rho(\omega', c')\, c'\, d\omega'\, dc' \tag{45}$$

or, using bracket notation

$$P\rho = c\langle \rho, c \rangle \tag{46}$$

The evolution of the system in the P subspace determines how the color

variables behaves. The first step in our analysis is therefore to describe the dynamics in the P subspace. For example, we must find a representation of the action of U in the P subspace. Using the definition of P above, and the description of U in terms of the symbolic dynamics [Eq. (44)], we find

$$PUP\rho = PUc\langle \rho, c\rangle \tag{47}$$

$$= P(-1)^{\chi(\omega)}\, c\langle \rho, c\rangle \tag{48}$$

$$= c\langle (-1)^{\chi(\omega)}\, c\rangle\langle \rho, c\rangle \tag{49}$$

$$= c(1-2p)\langle \rho, c\rangle \tag{50}$$

$$= (1-2p)\, P\rho \tag{51}$$

where $p = 2^{-k}$ is the area of the color-changing patch in the unit square. This means that we can write PUP as $(1-2p)\, P$.

A central object in the theory of subdynamics (see Section 3) is the resolvent operator $R(z)$. In the appendix, we use the methods of probability theory to find the P subspace representation $PR(z)\, P$ of the resolvent:

$$PR(z)\, P = \frac{(2z)^{1-k} + 2z^2 + 2^{1-k}(1-2z) - z - 1}{[(2z)^{1-k} + (z-1)(2z+1)](z-1)} \tag{52}$$

We also show, by considering the poles of $PR(z)\, P$, that the long-time behavior in the P subspace is exponentially decaying, with a certain decay constant λ. In fact, as $t \to \infty$, we have

$$\text{weak} \lim_{t \to \infty} PU^t P = P\beta\lambda^t \tag{53}$$

where β and λ are constants, and $|\lambda| < 1$. In the language of subdynamics, this allows us to write down an asymptotic P subspace resolvent (see Section 3),

$$P\Sigma(z)\, P = \frac{\beta}{z-\lambda} \tag{54}$$

Also, from Eq. (18), which we rearrange to give

$$\tilde{\psi}(z) = Pz - PUP - \frac{1}{PR(z)\, P} \tag{55}$$

we can obtain an expression for the collision operator $\tilde{\psi}(z)$,

$$\tilde{\psi}(z) = z + 2^{1-k} - 1 - \frac{(z-1)[(2z)^{1-k} + 2z^2 - z - 1]}{(2z)^{1-k} + 2^{1-k}(1-2z) + 2z^2 - z - 1} \tag{56}$$

These results demonstrate both the similarities and the differences between the urn and the pastry-cook model of Penrose and Coveney.[10] The pastry-cook model has long-time behavior of the same type as the urn, obeying an equation like (53). But the collision operator for the pastry-cook model is identically zero[10] and this makes the P subspace dynamics much simpler than for the urn. As one can see from Eq. (18), when $\tilde{\psi}(z) = 0$, the P subspace resolvent has only one pole, at $Pz = PUP$, or $z = 1 - 2p$. This means that $PR(z)\,P = P\Sigma(z)\,P$: the long-time behavior is no simpler than the general P subspace behavior. There is no separation into "subdynamics" in the P subspace. By contrast, in the baker's urn, the long-time limit allows a great simplification.

In work with the subdynamics formalism, use has often been made[4,17] of a "dissipativity condition." This condition guarantees that the amplitude of any mode in the P subspace decays with time rather than growing exponentially or remaining constant.

In the present case, exponential growth is not possible, because the evolution operator is unitary on the Hilbert space $\mathscr{H}$ of square-integrable phase space functions, and the P subspace is a subset of $\mathscr{H}$. Therefore to guarantee exponential decay, we need only confirm that $PR(z)\,P$ has no poles on the unit circle. A dissipativity condition which ensures this, analogous to the one formulated by Coveney and Penrose[17] for continuous-time systems, is the requirement

$$\|\tilde{\psi}(z) + PUP\| < 1 \qquad \text{when} \quad |z| = 1 \tag{57}$$

where $\|\cdot\|$ is the usual norm for the Banach space of linear operators from $\mathscr{H}$ onto itself, induced by the Hilbert-space norm $\|\cdot\|_{\mathscr{H}}$:

$$\|T\| := \sup_{\|x\|_{\mathscr{H}} = 1} \|Tx\|_{\mathscr{H}} \tag{58}$$

Such a condition is not necessary for our analysis of the baker's urn, since in the appendix we show that the largest pole in $PR(z)\,P$ has magnitude less than 1. However, it is not difficult to verify numerically that the requirement (57) is indeed satisfied.

In statistical mechanics, the most important purpose of the subdynamics formalism is the derivation of kinetic equations. Although in a discrete-time system it is not possible to find differential equations for the evolution of parts of the probability distribution, there is a discrete-time difference equation which is closely analogous to the general continuous-time kinetic equation which arises in subdynamics.

To derive this difference equation, we start with the equation for the time evolution of a single independent mode in the subdynamics:

$$\Pi\rho_{t+1} = U\Pi\rho_t \tag{59}$$

Now we operate on the left with the projector P and use the identity $1 = P + Q$:

$$P\Pi\rho_{t+1} = PUP\Pi\rho_t + PUQ\Pi\rho_t \tag{60}$$

Since, by Eq. (12), $Q\Pi = C(\lambda)\, P\Pi$, we have

$$P\Pi\rho_{t+1} = \Theta P\Pi\rho_t \tag{61}$$

where $\Theta = PU(C(\lambda) + P)\, P$. This is a closed kinetic equation for the P component of a single subdynamics mode. In the case where P is infinite-dimensional, such equations may be extremely complex. However, when P and therefore Π are finite-dimensional, the equation has a very simple form, as can be easily demonstrated using the results of Section 4. Since, by Eq. (31),

$$C(\lambda) + P = \tilde{\Pi} P[P\tilde{\Pi}P]^{-1} \tag{62}$$

and $U\tilde{\Pi} = \lambda\tilde{\Pi}$, we have $\Theta = \lambda$, and the kinetic equation is simply

$$P\Pi\rho_{t+1} = \lambda P\Pi\rho_t \tag{63}$$

This is the subdynamics kinetic equation for the baker's urn, and indeed for all systems analyzed using subdynamics with finite-dimensional projection operators.

7. ASYMPTOTIC OPERATORS FOR THE BAKER'S URN

If we wish to apply the asymptotic operator introduced in Section 2 to the baker's urn system, we must determine how the integrability condition discussed there restricts the space of possible ensemble densities ρ. Since a function which is continuous almost everywhere is Lebesgue-integrable, a matrix element $\langle \phi, U^t \rho \rangle$ will be well-defined when $\phi\tilde{\Pi}\rho$ is continuous except on a set of zero measure.

The condition of continuity on a function $f(x, y)$ can be expressed using the symbolic dynamics. Define the functional

$$\delta_n[f] := \max_{\omega} \left[\max_{\delta\omega:\, \delta\omega_i = 0,\, i < n} |f(\omega + \delta\omega) - f(\omega)| \right] \tag{64}$$

where the maximum is over all biinfinite strings of bits $\delta\omega_i \in \{-1, 0, 1\}$ which have all bits to the left of the nth place equal to zero.

A theorem that allows us to give a simple condition on the phase-space functions ϕ and ρ which ensures that $\langle \phi, \tilde{\Pi}\rho \rangle$ is well-defined is the following:

Theorem. If, for some finite α, $\delta_n[\rho]/\lambda^n \to \alpha$ as $n \to \infty$, then for all $y \in [0, 1)$, $\tilde{\Pi}\rho(x, y)$ is almost everywhere (a.e.) continuous as a function of x.

Proof. To prove this theorem, we need the following result.

Lemma. $\delta_n[\rho] \to 0$ as $n \to \infty$; then $\rho(x, y)$ is almost everywhere continuous as a function of x.

To see this, note that $\delta_n[\rho] \to 0$ means that for any $\varepsilon > 0$, we can find N_ε such that $n > N_\varepsilon$ implies $\delta_n[\rho] < \varepsilon$. Now for any point x which has a nonterminating binary expansion, choose

$$t_\varepsilon = \min_{i \in \mathbf{Z}} |x - 2^{-N_\varepsilon} i| \tag{65}$$

Then if $|\delta x| < t_\varepsilon$, the $\delta\omega$ equivalent to δx has no nonzero bits to the left of the N_εth, and so the $|\delta\rho|$ associated with δx is less than ε. Hence $\rho(x, y)$ is continuous at all points with a nonterminating binary expansion. Since the set of such points has measure one, ρ is continuous a.e. ∎

Now we proceed to the proof of the theorem. The lemma tells us that the function $f = \tilde{\Pi}\rho$ is continuous a.e. if

$$\begin{aligned} 0 &= \lim_{n \to \infty} \delta_n[f] \\ &= \lim_{n \to \infty} \lim_{r \to \infty} \lambda^{-r} \delta_n[U^r \rho_\perp] \end{aligned} \tag{66}$$

where $\rho_\perp = \rho - P^n\rho$ is the component of ρ orthogonal to the equilibrium distribution π. Hence, since $\delta_n[U\rho] = \delta_{n+1}[\rho]$, f is continuous if

$$0 = \lim_{n \to \infty} \lim_{r \to \infty} \lambda^{-r} \delta_{n+r}[\rho_\perp] \tag{67}$$

$$= \lim_{n \to \infty} \lim_{r \to \infty} \lambda^n \lambda^{-(n+r)} \delta_{n+r}[\rho_\perp] \tag{68}$$

$$= \lim_{n \to \infty} \lim_{s \to \infty} \lambda^n \lambda^{-s} \delta_s[\rho_\perp] \tag{69}$$

But if $\delta_n[\rho_\perp]/\lambda^n \to \alpha$, then the last line of (69) is equal to zero, and hence $f = \tilde{\Pi}\rho$ is continuous a.e. This proves the theorem. ∎

If we define the domain of $\tilde{\Pi}$ as the set of functions ρ which are integrable and satisfy the condition

$$\delta_n[\rho]/\lambda^n \to \alpha \tag{70}$$

for some finite α, then for bounded functions ϕ, $\langle \phi, \tilde{\Pi}\rho \rangle$ will be well-defined.

The condition (70) is a smoothness condition on ρ as a function of x, whose strength depends on the value of λ. For values of λ between 1/2 and 1, (70) is implied by uniform Lipschitz continuity with exponent $-\log_2 \lambda$. We may therefore take our space of allowable probability distributions to be the functions $\rho(x, y)$ which satisfy this Lipschitz continuity condition with respect to their x dependence and which are integrable. This ensures the convergence of $\langle \phi, \tilde{\Pi}\rho \rangle$ provided ϕ is bounded. A similar Lipschitz condition, derived in a different way, was used by Coveney and Penrose in their analysis of the pastry-cook model.[10]

This is one consistent way of ensuring convergence, but it is not the only one. The Lipschitz continuity condition can be regarded as a constraint which forces correlations of the function $\rho(x, y)$ to decay exponentially as it evolves forwards in time. We might call this "thermodynamic" behavior, since it reflects the tendency toward equilibrium which we see in physics. However, we could just as well ensure the convergence of matrix elements of $\tilde{\Pi}$ by allowing ρ to be any bounded function, and restricting ϕ to be Lipschitz continuous with respect to y. In this case the function ϕ has the same "thermodynamic" properties as ρ did in the first case, but with the direction of time reversed.

This choice between two solutions which are in a sense adjoint to one another and which are associated with different directions of time also arises in the rigged Hilbert spaces used by Antoniou and Tasaki[9] and Hasegawa and Saphir.[8] It is important to realize that these non-self-dual function spaces are not merely a mathematical convenience; they have direct physical implications. It is at this point that time asymmetry enters a previously time-symmetric theory.

It was shown by Penrose and Coveney[10] that for functions ϕ, ρ which are Lipschitz continuous, a matrix element of the operator $\tilde{\Pi}$ for the pastry-cook transformation can be written as

$$\langle \phi, \tilde{\Pi}\rho \rangle = \langle \phi, B \rangle \langle A, \rho \rangle \tag{71}$$

where A and B are distributions. A similar result holds for the baker's urn model, with the distributions A and B being defined by

$$\langle A, \rho \rangle = \lim_{r \to \infty} \lambda^{-r} \langle c, \rho(x)\, s_0 s_{-1} \cdots s_{r-1} \rangle \tag{72}$$

and

$$\langle \phi, B \rangle = \beta \lim_{r \to \infty} \lambda^{-r} \langle c, \phi(x) s_1 s_2 \cdots s_r \rangle \tag{73}$$

where $s_i = (-1)^{\chi(T^{1-i}\omega)}$. This result can be proved in the same manner as the corresponding result for the pastry-cook model. Here we justify Eq. (71) in a rather less rigorous way.

First consider

$$\lim_{t \to \infty} \langle \phi, U^t \psi \rangle \tag{74}$$

The components of ϕ and ψ which are even and odd in c will evolve separately under U, so we shall write them separately:

$$\phi(\omega, c) = \phi_e(\omega) + c\phi_o(\omega) \tag{75}$$

$$\psi(\omega, c) = \psi_e(\omega) + c\psi_o(\omega) \tag{76}$$

We then have

$$\begin{aligned} \langle \phi, U^t \psi \rangle = {} & \langle \phi_e, U^t \psi_e \rangle + \langle \phi_o, U^t c \psi_o \rangle \\ & + \langle \phi_e, U^t \psi_o \rangle + \langle c\phi_o, U^t \psi_e \rangle \end{aligned} \tag{77}$$

The third and fourth terms are zero. Correlation functions for the naked baker transformation decay at least as fast as 2^{-t}, as demonstrated by, for example, Hasegawa and Saphir.[8] Thus, in the limit $r \to \infty$, the first term cancels with the term $\langle \phi, P^\pi \psi \rangle$ in the matrix element of $\tilde{\Pi}$. This leaves the second term, which we can write using the symbolic dynamics as

$$\langle \phi_o(\omega), s_0 s_1 \cdots s_{t/2} s_{(t/2)+1} \cdots s_{t-1} \psi_o(T^{1-t}\omega) \rangle \tag{78}$$

This form is useful because Eq. (53) can be written using this notation as

$$\lim_{r \to \infty} \langle s_1 s_2 \cdots s_r, 1 \rangle = \beta \lambda^t \tag{79}$$

From (79), it follows that

$$\lim_{t \to \infty} \langle s_1 s_2 \cdots s_t, 1 \rangle = \beta \langle s_1 s_2 \cdots s_{t/2}, 1 \rangle \langle 1, s_{(t/2)+1} s_{(t/2)+2} \cdots s_t \rangle \tag{80}$$

Now in Eq. (78), as we take the limit $t \to \infty$, the correlation between $\phi_o(\omega)$ and $\psi_o(T^{1-t}\omega)$ decays, and correlations between the two parts of the

bracket depending on ω_i for $i > t/2$ and for $i \leqslant t/2$ are due solely to the s_i. Equation (78) then becomes

$$\langle \phi_o(\omega), s_0 s_1 \cdots s_{t/2} \rangle \beta \langle s_{(t/2)+1} \cdots s_{t-1}, \psi_o(T^{1-t}\omega) \rangle \tag{81}$$

and the matrix element for $\tilde{\Pi}$ is therefore

$$\langle \phi, \tilde{\Pi}\psi \rangle = \langle \phi, B \rangle \langle A, \psi \rangle \tag{82}$$

where A and B are the distributions defined by (72) and (73).

Equation (71) gives us a recipe for finding the asymptotic form of time-dependent correlation functions such as

$$\begin{aligned} g_{\phi\psi}(t) &= \langle \phi, U^t \psi \rangle - \langle \phi, 1 \rangle \langle 1, \psi \rangle \\ &= \langle \phi, U^t \psi \rangle - \langle \phi, P^\pi \psi \rangle \end{aligned} \tag{83}$$

since

$$\lim_{t \to \infty} g_{\psi\phi}(t) = \lambda^t \langle \phi, \tilde{\Pi}\psi \rangle \tag{84}$$

This last equation is correct provided that the correlation concerned has some component coupled to the color-decay mode, which is the slowest-decaying mode in the system. If the correlation has no such component, then it will decay faster than λ^t, and Eq. (84) will give the value zero for the correlation function.

It is interesting to compare our results with the work of Ruelle and others on resonances in dynamical systems.[12,13] These resonances are poles in the Fourier transforms

$$\tilde{g}_{\phi\psi}(\omega) = \sum_{t=-\infty}^{\infty} g_{\phi\psi}(t) z^{-t} \tag{85}$$

of correlation functions like Eq. (83). The positions of the poles do not depend on the functions ϕ and ψ, but the residue of a pole is given by $\langle \phi, \sigma_- \rangle \langle \sigma_+, \psi \rangle$, where σ_- and σ_+ are distributions (linear functionals).

The long-time behavior of a correlation function is then given by an equation of the form

$$\lim_{t \to \pm\infty} g_{\phi\psi}(t) = \langle \phi, \sigma_- \rangle \langle \sigma_+, \psi \rangle \alpha^t \tag{86}$$

where α is the location of the pole closest to the unit circle.

The pole λ which we have found looks similar to a Ruelle resonance, but one modification is necessary. To make the summation in (85) con-

verge, the correlation function (83) must decay backward as well as forward in time. This can be ensured by imposing the condition (70) for the limit $n \to -\infty$ as well as $n \to +\infty$. This imposes a continuity requirement on the y dependence as well as the x dependence of functions in the domain of $\tilde{\Pi}$. The pole at λ then gives a resonance at $z = \lambda$, and the residue of the pole is determined by $\langle \phi, B \rangle \langle A, \psi \rangle$ in Eq. (71).

The theory of resonances has been established rigorously only for Axiom A systems.[12,13] The baker's urn and pastry-cook's system do not satisfy Axiom A, because the maps which define them are not differentiable. However, the main results of the theory of resonances are derived using the fact that a symbolic dynamics can be constructed for Axiom A systems.[13] It is because a symbolic dynamics exists for each of the two model systems we have studied that the results of the theory of resonances apply to these systems.

8. DISCUSSION

The assumption that correlations decay exponentially with time is central to this paper, as it is to the subdynamics formalism[17] in the finite-dimensional cases studied here. If we wish to move from abstract dynamical systems into the "real world" of Hamiltonian dynamics, we must face the question of whether correlations in real systems decay exponentially. There are certainly systems of physical interest which do exhibit exponential decay. An example is the geodesic flow on a compact manifold of constant negative curvature.[23]

However, the existence of "long-time tails" in hydrodynamics[24] tells us that, in many cases, correlations do not decay exponentially. Also, Ruelle[25] has described a continuous-time dynamical system which shows that exponential decay of correlations is not a necessary condition for mixing. The rate of decay of correlations in "billiard" systems is still an open question. Only a subexponential bound on the decay has been established. Bunimovitch[26] has reviewed some significant results in the theory of model systems such as the Lorentz gas and billiards on variously shaped tables. In his opinion, exponential decay is likely to be "exotic" for such systems. However, Chernov[27] has established that decay is exponential in some dynamical systems which are closely related to billiards, and conjectures that his results can be generalized to billiard systems.

In our approach, there is the possibility of altering the asymptotic operator to extract a component of the dynamics which decays algebraically rather than exponentially. Since the subdynamics formalism relies heavily on the location of poles in matrix elements of the resolvent operator, it is not clear how it could be modified to cope with nonexponen-

tial decay, except when the projection operator P is infinite-dimensional. No rigorous results are known for this case.[17]

An important feature of the work we have presented is the way in which a restriction on the space of probability distributions arises naturally from the requirement that matrix elements of the operator $\tilde{\Pi}$ converge. This will apply to other systems as well as to the baker transformation. Many dynamical systems, including the geodesic flow mentioned above, can be represented by a symbolic dynamics which is a shift acting on some symbolic sequence. Continuity conditions like that given in Eq. (70) will arise if we apply the asymptotic operator to these systems.

Of course, deriving the space of probability distributions from the requirement of exponential decay in a sense amounts to putting the cart before the horse. Ideally, one should be able to deduce the space of distribution functions from some more fundamental principle which specifies the physically allowable ensembles. However, such a principle is lacking. The only widely discussed possibility is the "maximum entropy principle" of Bayesian statistics,[28,29] but this has been criticized on the grounds that the choice of prior is arbitrary.[30,31] These foundational issues have been reviewed by Penrose.[1]

APPENDIX. PROBABILITY THEORY AND THE DECAY OF COLOR IN THE BAKER'S URN

In this appendix, we consider the evolution of an ensemble density for the baker's urn which has the following probability distribution at time $t=0$:

$$\rho_0(x) = 1 + c(x) \tag{A1}$$

This distribution represents an ensemble which is normalized so that the total probability is 1, and which is uniform apart from a nonzero mean value of the color:

$$\langle c_0 \rangle = \langle c, \rho(x) \rangle = 1 \tag{A2}$$

The goal is to find the mean value of the color at time t, after t applications of the transformation:

$$\langle c_t \rangle = \langle c(x), U^t \rho(x) \rangle \tag{A3}$$

We can treat this problem using the theory of renewal processes,[32] because the time between the nth and $(n+1)$th color changes is a random variable whose distribution, after the first color change, is independent of n.

Let p_t be the probability that, given the starting distribution (A1), the first color change occurs at time t. Also let q_t be the probability that, if the nth color change occurs at time s, the $(n+1)$th occurs at time $s+t$. Then define the generating functions $G_1(z)$ and $G(z)$ by

$$G_1(z) := \sum_{t=1}^{\infty} \frac{p_t}{z^t} \tag{A4}$$

and

$$G(z) := \sum_{t=1}^{\infty} \frac{q_t}{z^t} \tag{A5}$$

Let $q_t^{(n)}$ be the probability that the nth change is at time t. The generating function for this quantity is

$$G_n(z) := \sum_{t=n}^{\infty} \frac{q_t^{(n)}}{z^t} = G_1(z)[G(z)]^{n-1} \tag{A6}$$

Therefore the probability that the nth change is at time t or before is

$$b_t^{(n)} = \sum_{r=n}^{t} q_r^{(n)} \tag{A7}$$

and the corresponding generating function is

$$H^{(n)} := \sum_{t=n}^{\infty} b_t^{(n)} z^{-t} \tag{A8}$$

$$= \sum_{t=n}^{\infty} \sum_{r=n}^{t} z^{-t} q_r^{(n)}$$

$$= \sum_{k=0}^{\infty} \sum_{r=n}^{\infty} z^{-(r+k)} q_r^{(n)} \tag{A9}$$

$$= \frac{1}{1-1/z} G_1(z)[G(z)]^{n-1} \tag{A10}$$

The probability that there will have been exactly n changes after time t can be written as

$$P(n, t) = b_t^n - b_t^{n+1} \tag{A11}$$

and so the mean value of the color at time t is given by

$$\langle c_t \rangle = \sum_{n=0}^{\infty} (-1)^n P(n, t) \tag{A12}$$

$$= 1 + 2 \sum_{n=0}^{\infty} (-1)^n b_t^{(n)} \tag{A13}$$

Finally, we define the color resolvent $R_c(z)$ by

$$R_c(z) := \frac{1}{z} \sum_{t=0}^{\infty} \langle c_t \rangle z^{-t} \tag{A14}$$

$$= \frac{1}{z} \sum_{t=0}^{\infty} z^{-t} + \frac{2}{z} \sum_{t=0}^{\infty} \sum_{r=1}^{\infty} (-1)^r b_t^{(r)} z^{-t} \tag{A15}$$

$$= \frac{1}{z-1} + \frac{2}{z} \sum_{r=1}^{\infty} (-1)^r H^{(r)}(z) \tag{A16}$$

Assuming the definition of P given in Section 6, the color resolvent $R_c(z)$ is the P subspace representation of the Brussels resolvent $R(z)$:

$$R_c(z) = PR(z)\, P \tag{A17}$$

Using (A10) and summing the series, we find

$$R_c(z) = \frac{2G_1(z) - G(z) - 1}{(1-z)[1 + G(z)]} \tag{A18}$$

To describe the behavior of $\langle c_t \rangle$ in terms of the resolvent $R_c(z)$, we must find the generating functions $G_1(z)$ and $G(z)$. To this end, we introduce two further generating functions:

$$F_1(z) := \sum_{t=1}^{\infty} u_t z^{-t} \tag{A19}$$

where u_t is the probability that any change of color will occur at time t, and

$$F(z) := \sum_{t=1}^{\infty} v_t z^{-t} \tag{A20}$$

where v_t is the probability that any color change will occur at time $r+t$, given that a color change occurred at time r. From the theory of renewal processes[32] we have

$$F_1(z)=\frac{G_1(z)}{1-G(z)} \tag{A21}$$

and

$$F(z)=\frac{G(z)}{1-G(z)} \tag{A22}$$

Combining these equations with (A18), we finally obtain

$$R_c(z)=\frac{2[F_1(z)-F(z)]-1}{[1+2F_1(z)](1-z)} \tag{A23}$$

The functions $F_1(z)$ and $F(z)$ can be calculated without difficulty. Since the probability distribution (87) is independent of ω, the probability of a change of color at time t is independent of t, being equal to the area of the color changing patch; thus $u_t=2^{-k}$ and so

$$F_1(z)=\frac{1}{2^k(z-1)} \tag{A24}$$

If there is a color change at time 0, the starting probability distribution is uniform in ω apart from the k binary digits $\omega_0\cdots\omega_{k-1}$, which are set equal to 1 in every member of the ensemble. These bits are shifted along t places to the left after time t, so the probability of a color change at time t is

$$v_t=\begin{cases}2^{-t} & \text{if } t\leqslant k\\ 2^{-k} & \text{if } t>k\end{cases} \tag{A25}$$

The generating function is therefore

$$F(z)=\frac{1-(2z)^{1-k}}{2z-1}+\frac{z}{(2z)^k(z-1)} \tag{A26}$$

and for the resolvent we have

$$R_c(z)=\frac{(2z)^{1-k}+2z^2+2^{1-k}(1-2z)-z-1}{[(2z)^{1-k}+(z-1)(2z+1)](z-1)} \tag{A27}$$

which is Eq. (52) in the main text.

We are interested in the asymptotic behavior of the system for large t. To find an asymptotic form for $R_c(z)$, we must locate the pole in $R_c(z)$ which is closest to the unit circle. This gives the mode in the color evolution which decays with the longest lifetime.

From Eq. (A27) it appears that $R_c(z)$ has a pole at $z=1$, but in fact the numerator has a factor $(z-1)$ and $R_c(z)$ is regular on the unit circle. By expanding the denominator in $p=2^{-k}$, the area of the color changing patch, we can find a pole which is on the real axis and close to the unit circle. To first order in p, the location $z=\lambda$ of the pole is

$$\lambda = 1 - \tfrac{2}{3}p \tag{A28}$$

We now show that, at least for $k \geqslant 6$, this is indeed the pole in Eq. (A27) closest to the unit circle. First, note that since the evolution operator U for our system is unitary, there can be no poles outside the unit circle. Let

$$g(z) = (2z)^{1-k} + (2z+1)(z-1) \tag{A29}$$

This is the denominator of $R_c(z)$. It is sufficient to show that for $k \geqslant 6$, there exists an $r_0 < \lambda$ for which there is exactly one zero z_0 of g with $|z_0| > r_0$.

Let $f(z) = (2z+1)(z-1)$, and denote the circle of radius r, centered on the origin, by C_r. We employ Rouché's theorem, using the functions f and g and the contour C_r. We have $|f-g| < |f|$ on C_r provided that $1 \leqslant r < 1/2$ and $(2r-1)(1-r) > (2r)^{1-k}$. A simple numerical calculation shows that this condition is satisfied when r takes the value $r_0 = 0.87$, provided $k \geqslant 6$. Rouché's theorem then tells us that for $k \geqslant 6$, there is exactly one zero of g outside the circle C_r. Since λ is outside C_r, this zero must be λ. Hence λ is the pole of R_c closest to the unit circle.

The asymptotic form of the resolvent is therefore

$$\Sigma_c(z) = \frac{\beta}{z-\lambda} \tag{A30}$$

where

$$\begin{aligned} \beta &= \lim_{z \to \lambda} (z-\lambda)\, R_c(z) \\ &= 1 - 4p/3 + O(p^2) \end{aligned} \tag{A31}$$

and so the asymptotic behavior of the color variable is given by

$$\lim_{t \to \infty} \langle c_t \rangle = \beta \lambda^t \tag{A32}$$

The formulas (A27), (A30), and (A32) are the final results of our probability-theoretic analysis. Although they give only a limited description of the system, they can be used as the starting point for the more general approach described in Section 7.

ACKNOWLEDGMENTS

P.V.C. is indebted to Oliver Penrose for many stimulating ideas and discussions which led to the present paper. We would like to wish him well on the occasion of his 65th birthday. We are also grateful to David Ruelle, Nikolai Chernov, and John Billingham for helpful comments. A.K.E. thanks the SERC for the award of a Ph.D. studentship.

REFERENCES

1. O. Penrose, Foundations of statistical mechanics, *Rep. Prog. Phys.* **42**:1937–2006 (1979).
2. P. C. W. Davies, *The Physics of Time Asymmetry* (Surrey University Press, 1974).
3. I. Prigogine, *Non-Equilibrium Statistical Mechanics* (Wiley, 1962).
4. R. Balescu, *Equilibrium and Non-Equilibrium Statistical Mechanics* (Wiley, New York, 1975).
5. I. Prigogine, C. George, and F. Henin, Dynamical and statistical descriptions of N-body systems, *Physica* **45**:418–434 (1969); I. Prigogine, C. George, F. Henin, and L. Rosenfeld, A unified formulation of dynamics and thermodynamics, *Chem. Scripta* **4**:5–32 (1973).
6. T. Y. Petrosky and H. Hasegawa, Subdynamics and non-integrable systems, *Physica A* **160**:351–385 (1989).
7. A. P. Grecos and I. Prigogine, Kinetic and ergodic properties of quantum systems—The Friedrichs model, *Physica* **59**:77–96 (1972); I. E. Antoniou and I. Prigogine, Intrinsic irreversibility and integrability of dynamics, *Physica A* **192**:443–464 (1993).
8. H. H. Hasegawa and W. C. Saphir, Non-equilibrium statistical mechanics of the baker map: Ruelle resonances and subdynamics, *Phys. Lett. A* **161**:477–482 (1992); Unitarity and irreversibility in chaotic systems, *Phys. Rev. A* **46**:7401–7423 (1992).
9. I. E. Antoniou and S. Tasaki, Generalized spectral decompositions of mixing dynamical systems, *Int. J. Quant. Chem.* **46**:425–474 (1993).
10. O. Penrose and P. V. Coveney, Is there a canonical nonequilibrium ensemble? preprint (1994).
11. P. V. Coveney and A. K. Evans, Long time evolution in statistical mechanics: subdynamics, dynamical resonances and canonical nonequilibrium ensembles, preprint (1994).
12. D. Ruelle, Resonances of chaotic dynamical systems, *Phys. Rev. Lett.* **56**:405–407 (1986).
13. J. P. Eckmann, Resonances in dynamical systems, in *Proceedings of the IXth International Congress on Mathematical Physics* (Institute of Physics Publishing, Bristol, 1989), pp. 192–207.
14. F. Christiansen, G. Paladin, and H. H. Rugh, Determination of correlation spectra in chaotic systems, *Phys. Rev. Lett.* **65**:2087–2090 (1990).
15. A. Lasota and M. C. Mackey, *Probabilistic Properties of Deterministic Systems* (Cambridge University Press, Cambridge, 1985).

16. V. I. Arnold and A. Avez, *Ergodic Problems of Classical Mechanics* (Benjamin, New York, 1968).
17. P. V. Coveney and O. Penrose, On the validity of the Brussels formalism in statistical mechanics, *J. Phys. A* **25**:4947–4966 (1992).
18. O. F. Bandtlow and P. V. Coveney, On the discrete-time version of the Brussels formalism, preprint (1994); O. F. Bandtlow, On the discrete time version of the Brussels formalism, M. Phil. Thesis, Cavendish Laboratory, University of Cambridge (1992).
19. P. V. Coveney and C. George, On the time-dependent formulation of analytical continuation in statistical mechanics, *Physica A* **141**:403–426 (1987); P. V. Coveney, Subdynamics in the time-dependent formalism, *Physica A* **113**:123–146 (1987).
20. A. Böhm, *Quantum Mechanics* (Springer-Verlag, 1979).
21. D. Ruelle, *Elements of Differentiable Dynamics and Bifurcation Theory* (Academic Press, 1989).
22. R. Jancel, *Foundations of Classical and Quantum Statistical Mechanics*, Appendix III (Pergamon, 1969).
23. T. Bedford, M. Keane, and C. Series, (eds.), *Ergodic Theory, Symbolic Dynamics and Hyperbolic Spaces* (Oxford University Press, Oxford, 1991).
24. L. E. Reichl, *A Modern Course in Statistical Mechanics* (Arnold, 1980).
25. D. Ruelle, Flots qui ne mélangent pas exponentiellement, *C. R. Acad. Sci. Paris Ser. I* **296**:191–193 (1983).
26. L. A. Bunimovitch, Decay of correlations in dynamical systems with chaotic behaviour, *Sov. Phys. JETP* **62**:842–852 (1985).
27. N. I. Chernov, Ergodic and statistical properties of piecewise linear hyperbolic automorphisms of the 2-torus, *J. Stat. Phys.* **69**:111–136 (1992).
28. W. T. Grandy, *Foundations of Statistical Mechanics* (Reidel, Dordrecht, 1987).
29. E. T. Jaynes, *Papers on Probability, Statistics and Statistical Mechanics*, R. D. Rosenkrantz, ed. (Reidel, Dordrecht, 1983).
30. J. M. Jauch and J. G. Baron, Entropy, information and Szilard's paradox, *Helv. Phys. Acta* **45**:220 (1972).
31. O. Penrose, *Foundations of Statistical Mechanics* (Pergamon, 1970).
32. W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. I, 2nd ed. (Wiley, New York, 1957).