

On the long-time behaviour of ensembles in a model of deterministic diffusion

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

1998 J. Phys. A: Math. Gen. 31 5887

(<http://iopscience.iop.org/0305-4470/31/28/006>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 171.66.16.122

The article was downloaded on 02/06/2010 at 06:57

Please note that [terms and conditions apply](#).

On the long-time behaviour of ensembles in a model of deterministic diffusion

Allan K Evans[†] and Peter V Coveney[‡]

[†] Department of Mathematical Sciences, De Montfort University, The Gateway, Leicester LE1 9BH, UK

[‡] Schlumberger Cambridge Research, High Cross, Cambridge CB3 0EL, UK

Received 6 February 1998

Abstract. We describe a large model system, based on the baker transformation, where deterministic diffusion occurs. The model is similar to one recently considered by Gaspard, and by Hasegawa and Driebe. We point out the close relationship between this system and a simple random walk, and analyse the evolution with time of ensembles in the system using the resolvent-based version of the ‘subdynamics’ formalism developed by Prigogine and his collaborators. We obtain an exact and rigorous description of the long-time behaviour of ensembles, including the irreversible approach to equilibrium, for the case where the system has finite size. We also consider the ‘thermodynamic’ limit where the size of the system becomes infinite, and derive a description of the long-time behaviour in this case, where correlations decay non-exponentially with time.

1. Introduction

The subject of this paper is the process by which diffusion arises from microscopic chaos. Our model is a discrete-time dynamical system, based on the well known baker transformation (Arnold and Avez 1968). The model is deterministic and reversible, but we show that a certain class of ensembles approaches equilibrium irreversibly. The model system is spatially extended, consisting of N cells arranged in a line. For finite N , the decay of correlations with time is exponential, but in the limit $N \rightarrow \infty$, the decay becomes non-exponential.

A similar model system was devised by Gaspard (1992). Using the theory of resonances in dynamical systems, he showed that the longest-lived resonances corresponded to the eigenvalues of a phenomenological diffusion equation. Both Gaspard (1993) and Hasegawa and Driebe (1994) gave a spectral decomposition of the evolution operator for Gaspard’s system.

Other authors have also considered the relationship between diffusion and chaos. A commonly used model, rather more physical than the one employed here, is the Lorentz gas. Garrido and Gallavotti (1994) carried out a recent numerical study of this system. Gaspard (1993) also considered a variant of the Lorentz gas, the ‘Lorentz channel’. Schuster (1988) treated a simple discrete-time model of deterministic diffusion, and Giesel and Nierwetberg (1982) found universal features of the onset of diffusive behaviour as a parameter is varied. We shall be concerned not with the onset of diffusion, but with the way in which apparently random large-scale behaviour emerges from chaotic dynamics on a small scale.

In this paper, after introducing the dynamical system, we show that it is closely related to the stochastic process generated by a simple Markovian random walk on the integers. The link between the deterministic dynamical system and the stochastic process of the random walk is a central theme of this paper. The link is strengthened in section 3.2, when we derive a kinetic equation. This kinetic equation is the same as the evolution equation for the random walk, and is obeyed exactly at long times by a certain reduced description of ensembles in our model system.

The main part of this paper is concerned with applying the ‘subdynamics’, or ‘Brussels’ formalism developed by Prigogine and his collaborators. Using this formalism, we describe the long-time behaviour of ensembles in our model system, for finite N . The use of symbolic dynamics simplifies the work. To make mathematical sense, the subdynamics theory must be placed in a well-defined space of ensembles. Certain steps in the application of the theory can be justified only if the linear operators involved (such as the time evolution operator) satisfy mathematical conditions formulated in this space. As we go through the process of applying the theory in section 3, we note the conditions required to justify these steps. In section 4, we construct a space of ensembles in which the required conditions hold, and hence show that the ‘subdynamics’ description of the long-time behaviour of ensembles is correct, at least for ensembles contained in the space.

In all discrete-time systems to which the subdynamics formalism has been applied previously, the decay of ensembles is exponential at long times. This is the case where the application of the subdynamics formalism is most straightforward. The spectrum of the evolution operator is discrete, and each exponentially decaying mode in the system corresponds to a single simple pole in the resolvent. For finite values of N , this is also true for the model system we consider here. However, in the limit $N \rightarrow \infty$, the discrete spectrum becomes continuous and the decay of correlations becomes non-exponential at long times. As far as the authors know, this is the first case where the subdynamics formalism has been applied to a discrete-time system where correlations decay non-exponentially with time.

Some previous work on the application of the theory of subdynamics, and related methods, to discrete-time dynamical systems has been done by Hasegawa and Saphir (1992), Antoniou and Tasaki (1993), and Hasegawa and Driebe (1994). The approach taken by these authors is to work with matrix elements of operators in a structure known as a rigged Hilbert space. Here, as in our previous paper (Evans and Coveney 1995), rather than working with matrix elements we consider bounded linear operators on a Banach space of ensembles.

2. The deterministic random walk

2.1. The dynamical system

We now describe the simple model of deterministic diffusion which is the subject of this paper. The model is a discrete-time dynamical system, defined by a phase space Γ , and a map T from Γ onto itself. T carries phase-space points forwards in time: if $x \in \Gamma$ is the state of the system at time t , then Tx is the state at time $t + 1$. For our system, the phase space is N copies of the unit square

$$\{(x, y) : 0 \leq x < 1, 0 \leq y < 1\}. \quad (1)$$

These copies are arranged in a row, and labelled with integers r running from 1 to N . We use periodic boundary conditions to simplify the calculations; however the boundary conditions will make no difference in the limit $N \rightarrow \infty$, because in this limit an initially

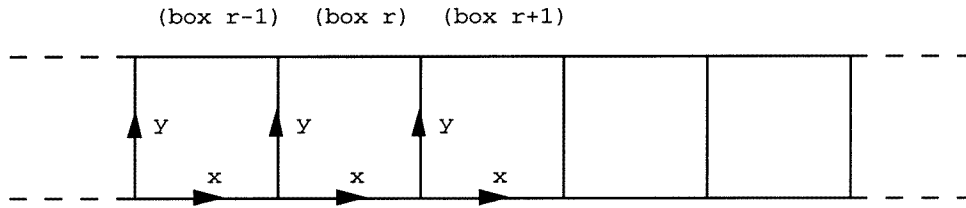


Figure 1. The phase space for the DRW.

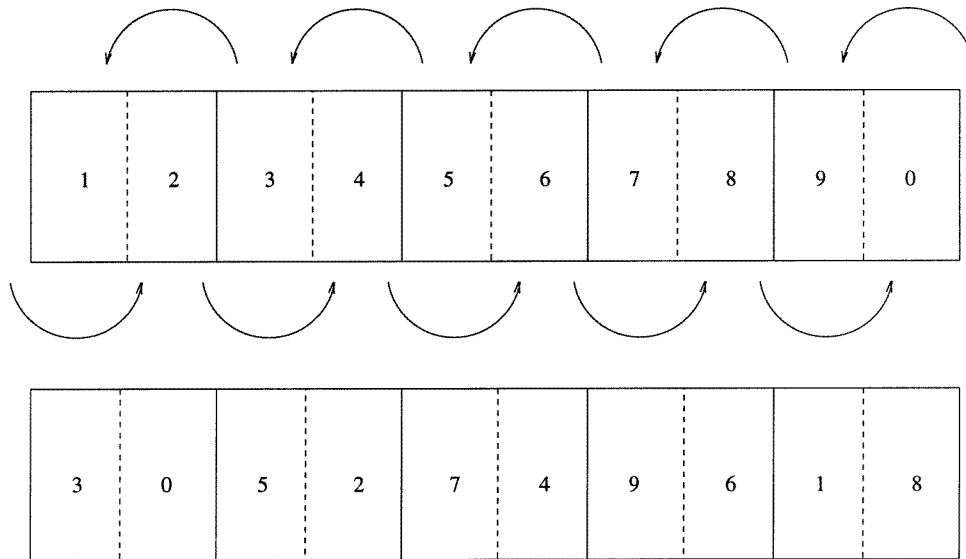


Figure 2. The first stage of the map T .

localized distribution never reaches the boundary.

A point in the phase space is specified by three variables (r, x, y) , with r labelling the square, and (x, y) giving a point in the square (figure 1).

The dynamical map T can be visualized as a geometrical operation on the phase space in two steps. The first step is to move the left-hand half of each square one space to the left, and the right-hand half one space to the right (figure 2).

The second step is to perform a baker transformation twice successively in each square. The baker transformation is the mapping

$$B(x, y) = \begin{cases} \left(2x, \frac{y}{2}\right) & \text{if } x < \frac{1}{2} \\ \left(2x - 1, \frac{y}{2} + \frac{1}{2}\right) & \text{if } x \geq \frac{1}{2}. \end{cases} \quad (2)$$

As a dynamical system, the baker transformation is highly chaotic. It belongs to the class of systems known in ergodic theory as Bernoulli automorphisms (Petersen 1983). Such systems are ergodic and mixing.

To visualize the baker 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, again forming a

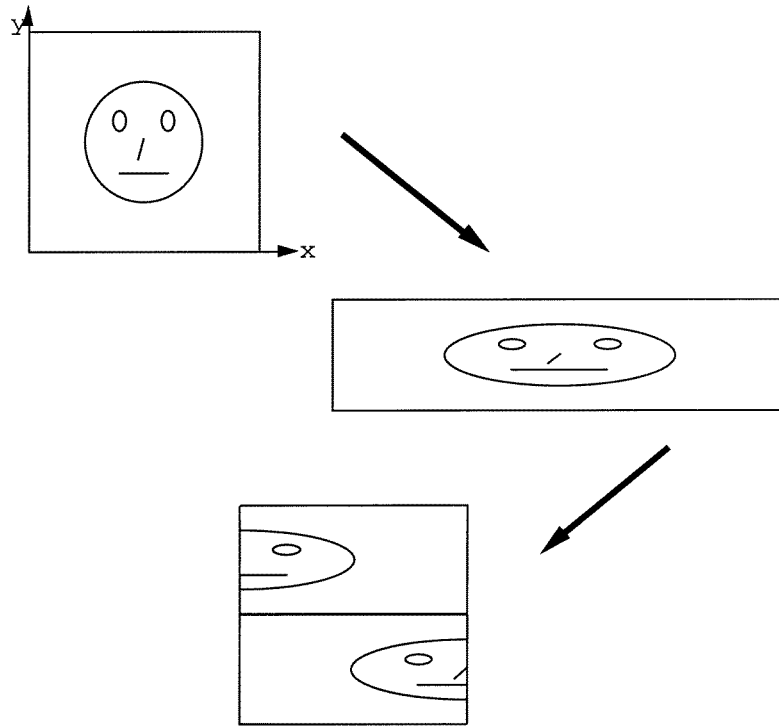


Figure 3. The baker transformation.

square (figure 3).

The baker transformation is easy to analyse because of its simple symbolic dynamics. Any point in the unit square S can be described by a bi-infinite string of binary digits $\dots\omega_{-2}\omega_{-1}\omega_0\omega_1\dots$. The infinite string $\omega_1\omega_2\omega_3\dots$ is the binary expansion of x , and $\omega_0\omega_{-1}\omega_{-2}\dots$ is y in binary. The baker transformation is then a shift of the string one place to the left:

$$(B\omega)_i = \omega_{i+1}. \quad (3)$$

If we write a single point in the row of unit squares as (r, ω) , where r specifies a square and ω a point in the square, then the map T is described by the equation

$$T(r, \omega) = (r - (-1)^{\omega_1}, B^2\omega). \quad (4)$$

We will refer to the dynamical system described by the map T as the deterministic random walk (DRW). The relationship between this system and the random walk is described in section 2.4.

In physical terms, one may think of the system as a crude model of an ideal gas, contained in a vessel which consists of a series of chambers, each linked to the next by a hole (figure 4). Penrose and Coveney (1994) gave a similar physical picture for their 'pastry-cook' transformation. Over long time periods and on large length scales, we intuitively expect that the details of the deterministic motion will be unimportant. A coarse-grained picture of the dynamics should look like a simple random walk between the boxes. More generally, the DRW is intended to represent any physical system where diffusive behaviour on a large scale results from microscopic deterministic chaos. This is common in nature: turbulence, the conduction of heat, and diffusion are examples.

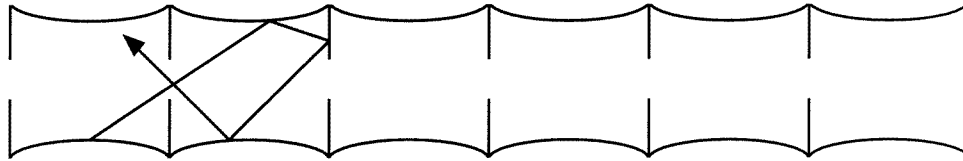


Figure 4. A physical picture of the DRW model.

Using two successive baker transformations, rather than just one, speeds the decay to equilibrium of non-uniformities in the (x, y) dependence of ensembles. This is necessary because when only one baker transformation is used, the diffusive behaviour which is the main subject of this paper does not always dominate the long-time behaviour of ensembles. The mixing of the (x, y) coordinates in the phase space is so slow that it may significantly influence the long-time behaviour. This point is discussed in section 4.5.

The idea of coupling strings of baker transformations together to build a model system of this kind is not new. Elskens and Kapral (1985) devised a system of three coupled baker transformations, and recently we have studied a system of two coupled baker transformations in collaboration with Penrose (Penrose and Coveney 1994, Evans and Coveney 1995). A slightly more complicated version of this system is discussed in Coveney and Evans (1994). The system that is the subject of this paper differs only in detail from the ‘multi-baker map’ of Gaspard (1992), both systems being constructed by linking chains of baker transformations. Gaspard’s work was mainly concerned with the theory of resonances in dynamical systems. Vollmer *et al* (1997) also analysed the effect of coarse-graining on the entropy production in a similar model, and considered the implications for non-equilibrium statistical mechanics.

2.2. Dynamics of ensembles

For the DRW, an ensemble is represented by a function $\rho(r, \omega)$. Physically, an ensemble should be a non-negative integrable function, but it will be convenient to work with a larger space which includes functions that may be negative, and some objects that are not functions but linear functionals (distributions). This space is defined in detail in section 4. We define an evolution operator U which carries ensembles forward in time in the same way that T carries points:

$$U\rho(x) = \rho(T^{-1}x). \tag{5}$$

For the DRW, we have

$$U\rho(r, \omega) = \rho(r + (-1)^{\omega-1}, B^{-2}\omega). \tag{6}$$

A Fourier transform representation of ρ is useful:

$$\rho_\theta(\omega) = \sum_r \rho(\omega, r)e^{-ir\theta}. \tag{7}$$

If N , the range of r , is finite, then the inverse transform is

$$\rho(\omega, r) = \frac{1}{N} \sum_{p=1}^N \rho_\theta(\omega)e^{i\theta_p r} \tag{8}$$

where $\theta_p = 2\pi p/N$. If we let $N \rightarrow \infty$, then we have

$$\rho(\omega, r) = \frac{1}{2\pi} \int_0^{2\pi} d\theta \rho_\theta(\omega)e^{i\theta r}. \tag{9}$$

We will assume that all density functions ρ are real, so that $\rho_{-\theta}(\omega) = \rho_{\theta}^*(\omega)$.

The action of U on this representation is

$$U\rho(\omega, r) = \frac{1}{N} \sum_{p=1}^N [e^{is_{-1}(\omega)\theta_p} \rho_{\theta_p}(B^{-2}\omega)] e^{ir\theta_p} \quad (10)$$

where

$$s_i(\omega) = (-1)^{\omega_i}. \quad (11)$$

2.3. The reduced description

The subdynamics theory employs a reduced description of the system, which is determined by the probability density $\rho(x)$ but carries less information. An example of such a description in statistical mechanics is the one-particle distribution function in the Boltzmann equation. The reduced description is specified by a projection operator P , satisfying $P^2 = P$. Its complement $Q = 1 - P$ is also a projector. We choose P so that information about the most slowly changing variables in the system is contained in $P\rho$. The procedure which we follow in section 3 then gives us the long-time behaviour of ensembles over the dynamical system's phase space.

It is not always obvious how to choose an appropriate projector P for a given dynamical system. Since the success of the subdynamics formalism depends on this choice, care is necessary. In Evans and Coveney (1995), we have given an example of how an inappropriate choice can lead to incorrect results. In section 4.5 we shall see that the same is true for the DRW.

For the DRW, our choice of reduced description is motivated by the analogy between the system and the diffusive processes mentioned in section 2.1. We choose a description which contains information only on the way particles are distributed between boxes, and not on the distribution within each box. We hope that this coarse-grained description will satisfy a simple kinetic equation over large timescales. This kinetic equation is derived in section 3.2.

For the DRW, we define the operator P , acting on an ensemble ρ , by

$$\begin{aligned} P\rho(x, y, r) &= \int dx dy \rho(x, y, r) \\ &= \langle \rho \rangle \end{aligned} \quad (12)$$

where $\langle \dots \rangle$ is used as shorthand for integration over the unit square. The function $p(r) = P\rho$ depends only on r , and not on x or y . For any r , $p(r)$ gives us the probability of finding a particle in the box labelled by r .

The integral can also be expressed in terms of the symbolic sequence ω :

$$\begin{aligned} \langle f \rangle &= \int dx dy f(x, y) \\ &= \lim_{n \rightarrow \infty} \frac{1}{2^{2n+1}} \sum_{\omega_{-n}=0,1} \dots \sum_{\omega_n=0,1} f(\omega). \end{aligned} \quad (13)$$

Using the symbolic description in terms of ω , and the Fourier transform representation of equation (7), we can find expressions for the operators PU^tP which give the action of U in the P -subspace. From equations (10) and (12), we have

$$UP\rho(\omega, r) = \frac{1}{N} \sum_{p=1}^N [e^{is_{-1}(\omega)\theta_p} \langle \rho_{\theta_p} \rangle] e^{ir\theta_p} \quad (14)$$

and

$$PUP\rho(\omega, r) = \frac{1}{N} \sum_{p=1}^N [\cos \theta_p \langle \rho_{\theta_p} \rangle] e^{ir\theta_p} \tag{15}$$

since

$$Pe^{is_{-1}(\omega)\theta_p} = \langle e^{is_{-1}(\omega)\theta_p} \rangle = \cos \theta_p. \tag{16}$$

In the same way, we can show that for $t > 0$,

$$PU^t P\rho(\omega, r) = \frac{1}{N} \sum_{p=1}^N [\cos^t \theta_p \langle \rho_{\theta_p} \rangle] e^{ir\theta_p} \tag{17}$$

so that $(PUP)^t = PU^t P$. This property of the projection operator P will greatly simplify future calculations.

2.4. The random walk

We now compare the P -subspace dynamics of the DRW with a simple Markovian random walk on the integers r , defined as follows.

Let an integer r be the state of the system at time t . Then the state at time $t + 1$ is determined by tossing a coin and adding 1 to r if the coin shows a head, and subtracting 1 otherwise. An ensemble is represented by a function $f(r)$. The evolution operator for this system is V , where

$$Vf(r) = \frac{1}{2}[f(r + 1) + f(r - 1)]. \tag{18}$$

By applying this operator to the Fourier representation (equation (7)), we can show that the action of V^t on a function $f(r)$ is the same as the action of $PU^t P$, as in equation (17). Hence, for any function $f(r)$, and for each positive integer t ,

$$PU^t Pf(r) = V^t f(r). \tag{19}$$

That is, the action of U in the P -subspace is the same as that of the evolution operator for a simple random walk.

Another way of stating this result is if we start at $t = 0$ with an ensemble $\rho(r)$ which is independent of ω , then the reduced description of the ensemble at time t , given by $PU^t \rho(r)$, is the same as if ρ had evolved according to the simple random walk. This is why we refer to the system described by T and U as a ‘deterministic random walk’. The dynamics of the DRW are completely deterministic, but the behaviour of the reduced description defined by P is the same as for a random walk. This also suggests that for infinite N , non-exponential decay will arise in the DRW in the same way that diffusive behaviour emerges from a random walk.

3. Subdynamics

In this section, we apply the subdynamics formalism to find the long-time behaviour of ensembles in the deterministic random walk. We shall not give details of the mathematical background to the formalism in this paper. The discrete-time version of the theory used here has also been used in Evans and Coveney (1995), and analysed in detail by Bandtlow and Coveney (1994). Balescu’s book (1975) contains a treatment of the continuous-time version of the formalism, with applications to statistical mechanics, as it stood in the mid 1970s. A more concise and critical account is given by Obcemea and Brändas (1983). A recent

formulation of the theory, not involving the resolvent, has been considered by Petrosky and Hasegawa (1989) and Antoniou and Tasaki (1993).

As we describe the subdynamics calculations, we shall note some mathematical assumptions that are necessary. These assumptions are bounds on linear operators such as the evolution operator U . In section 4, we construct a Banach space of ensembles in which these conditions are satisfied.

3.1. P -subspace operators

A central object in the theory is the resolvent operator $R(z) = 1/(z - U)$. If U is a bounded operator on a Banach space, then the resolvent, considered as a function on a subset of the complex plane taking values in the space of bounded operators, is holomorphic at infinity and has a Laurent series expansion given by

$$R(z) = \sum_{t=0}^{\infty} \frac{U^t}{z^{t+1}}. \quad (20)$$

This expansion is valid when $|z| > \|U\|$, where $\|U\|$ is the operator norm of U (Reed and Simon 1972). The boundedness of U is the first condition which must be satisfied by the space of ensembles defined in section 4.

In applying the subdynamics theory to a specific system, we first find an expression for the P -projection of the resolvent, $PR(z)P$. We think of $PR(z)P$, and other P -subspace operators such as $\psi(z)$, as operating not on the full set of possible ensembles, but on the P -subspace, which is $\text{Ran } P$, the range of P . When N is the number of cells in the DRW's phase space, the dimension of $\text{Ran } P$ is N . We can therefore write the P -subspace operators as $N \times N$ matrices operating on some basis of this finite-dimensional space. The question of which space of ensembles U acts upon does not influence this representation, so long as $\text{Ran } P$ is included in the space.

Each singularity of the operator $PR(z)P$ marks a decaying 'mode' in the P -subspace. If the singularities are simple poles z_i , as for the DRW when N is finite, then each mode decays exponentially with time, as z_i^t .

From equations (17) and (20), we find that

$$PR(z)P\rho(\omega, r) = \frac{1}{N} \sum_{p=1}^N e^{ir\theta_p} \left[\frac{1}{z - \cos \theta_p} \langle \rho_{\theta_p} \rangle \right]. \quad (21)$$

Since P is bounded, this equation is valid where equation (20) is valid; that is, for $|z| > \|U\|$. However, the right-hand side of equation (21) is a holomorphic function on the entire complex plane except at the poles $z_p = \cos \theta_p$. Thus equation (21) defines a meromorphic continuation of $PR(z)P$ from the region $|z| > \|U\|$ to the entire complex plane. This continuation is the meromorphic function referred to in proposition 1 of Bandtlow and Coveney (1994), and therefore gives a correct representation of $PR(z)P$ in the region $|z| > \|QUQ\|$.

We now wish to obtain an operator which is related to $PR(z)P$, but which takes into account only the most slowly decaying modes in the system's evolution. Since each pole of equation (21) at $z_p = \cos \theta_p$ gives rise to a mode which decays exponentially as $(\cos \theta_p)^t$, we retain the poles which have $|\cos \theta_p| > \lambda$, and discard the others, where λ is a threshold value. Provided we choose $\lambda > \|QUQ\|$, these poles are in the domain of validity of equation (21) and are therefore shared by $PR(z)P$. We thus obtain an operator $W(z)$,

defined by

$$W(z)\rho(\omega, r) = \frac{1}{N} \sum_{p: |\cos \theta_p| > \lambda} e^{ir\theta_p} \left[\frac{1}{z - \cos \theta_p} \langle \rho_{\theta_p} \rangle \right]. \tag{22}$$

This operator $W(z)$ is usually written as $P\Sigma(z)P$, because it is related by the equation

$$W(z) = P\Sigma(z)P \tag{23}$$

to an asymptotic operator $\Sigma(z)$ which we shall construct later. However, for the moment $W(z)$ should be thought of simply as an operator from the P -subspace into itself which contains information about long-time behaviour in the P -subspace.

The collision operator $\psi(z)$ is a measure of the influence of the Q -subspace on the P -subspace dynamics. It is defined by

$$\psi(z) = P U Q \frac{1}{z - Q U Q} Q U P. \tag{24}$$

Since (see Bandtlow and Coveney 1994),

$$P R(z) P = P \frac{1}{z - P U P - \psi(z)} P \tag{25}$$

we can show using equation (21) that $\psi(z) = 0$, this conclusion also being valid for $|z| > \|QUQ\|$.

As we stated earlier, these P -subspace operators can be defined independently of the function space used as the domain of U . An important stage of the subdynamics procedure is to extend the long-time description given by $W(z)$ from the P -subspace into the full function space. This is the subject of section 3.3.

3.2. The kinetic equation

In statistical mechanics, the most important purpose of the subdynamics formalism is the derivation of kinetic equations such as the Boltzmann equation (Balescu 1975). Although for 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 continuous-time kinetic equation derived using subdynamics.

In this case, where $\psi(z) = 0$, the equation is simply

$$\tilde{\rho}_{t+1} = P U P \tilde{\rho}_t \tag{26}$$

as we have shown in an earlier paper (Evans and Coveney 1995). This is a closed kinetic equation for the evolution of the P -subspace component of the mode that dominates the long-time behaviour. In the long-time limit, the reduced distribution

$$\tilde{\rho}_t(r) = \langle \rho_t(\omega, r) \rangle = P \rho_t(\omega, r) \tag{27}$$

converges to a solution of this equation. Since $V = P U P$, where V is the evolution operator for the random walk of section 2.4, in the long-time limit the reduced distribution function evolves according to

$$\tilde{\rho}_{t+1}(r) = V \tilde{\rho}_t(r) \tag{28}$$

as if it is governed by a random walk.

This result strengthens the connection between the DRW and the simple random walk. In section 2.4, we noted that for any ensemble $\rho(r)$ which is independent of ω at $t = 0$, the reduced description at time t is the same as if ρ had evolved according to the simple

random walk. Now we see that for a more general class of ensembles $\rho(r, \omega)$ at $t = 0$, at long times the reduced description converges to the behaviour of a simple random walk.

At this point, we have not constructed the mathematical framework of the theory, so we do not yet know for which ensembles $\rho(r, \omega)$ this conclusion is valid. This mathematical framework is the subject of section 4. There, we shall see that the theory is valid only when the number of cells N is 5 or greater than 6, and only for certain classes of ensembles ρ .

In the remaining part of section 3, we give a more complete description of the long-time behaviour of ensembles in the DRW.

3.3. Moving out of the P -subspace

Since the P -subspace was chosen so that it included the ‘slow variables’ of the system, we would expect that those modes included in $W(z)$ dominate the long-time behaviour, even outside the P -subspace. The next step is therefore to extend $W(z)$ to include the Q -subspace, giving an asymptotic operator $\Sigma(z)$ such that $W(z) = P\Sigma(z)P$. In section 4, we show that this operator does indeed give an accurate representation of the long-time behaviour of ensembles.

The equation which allows this extension to be carried out is the ‘Brussels decomposition’:

$$R(z) = [P + C(z)]PR(z)P[P + D(z)] + S(z). \quad (29)$$

For details of the derivation of this equation, see Bandtlow and Coveney (1994). The operators $S(z) = Q[1/(z - QUQ)]Q$, $C(z) = S(z)UP$ and $D(z) = PUS(z)$ are known as the reduced resolvent, the creation operator, and the destruction operator. These operators are holomorphic in the region $|z| > \|QUQ\|$. Bandtlow and Coveney showed that equation (29) is valid in this region of the complex plane.

Because $C(z)$, $D(z)$ and $S(z)$ are holomorphic in this region, (29) implies that the only singularities of $R(z)$ in the region $|z| > \|QUQ\|$ are simple poles at the location of the simple poles in $PR(z)P$. Each of these poles represents a decaying mode of the evolution described by U . We can therefore construct an asymptotic operator $\Sigma(z)$ using equation (29). We set $\lambda = \|QUQ\|$ and let

$$W_j = \lim_{z \rightarrow z_j} (z - z_j)PR(z)P \quad (30)$$

be the residue at each simple pole z_j of $W(z)$, and then construct $\Sigma(z)$ by

$$\Sigma(z) = \sum_{z_j} [P + C(z_j)] \frac{W_j}{(z - z_j)} [P + D(z_j)] \quad (31)$$

where the sum is over the poles of $W(z)$. The operator $\Sigma(z)$ then has exactly the same singularities as $R(z)$ in the region $|z| > \|QUQ\|$. We can find the time dependence of the modes described by $\Sigma(z)$ by performing the inverse to the transform operation of equation (20). We thus obtain a time-dependent asymptotic operator Σ_t :

$$\Sigma_t = \oint z^t \Sigma(z) dz \quad (32)$$

where the integral is around a contour in the complex plane enclosing the unit circle.

The operator Σ_t , gives an asymptotic description of the long-time behaviour of ensembles. In other words, for the class of ensembles $\rho(x, y, r)$ for which the theory is valid (see section 4),

$$\lim_{t \rightarrow \infty} \|U^t \rho - \Sigma_t \rho\| = 0 \quad (33)$$

where the limit is in the topology of the Banach space described in section 4. As we shall see, we can also show, for finite N , that the convergence of the limit in equation (33) is exponential. Setting $t = 0$ in the operator Σ_t yields the operator Π , a projection operator which, when applied to any ensemble, yields the component of that ensemble which controls the long-time behaviour. For more details, see Bandtlow and Coveney (1994) and Evans and Coveney (1995).

All the calculations in this section have depended on the condition that QUQ is bounded, and that $\|QUQ\|$ is small enough so that some of the singularities of $W(z)$ lie in the region $|z| > \|QUQ\|$. This is the second condition that will be considered in section 4.

3.4. Creation and destruction operators

We now use the Fourier representation of section 2.2 to find expressions for the creation and destruction operators $C(z)$ and $D(z)$. First, we consider the creation operator $C(z)$, which has a Laurent expansion

$$C(z) = \sum_{t=1}^{\infty} \left(\frac{QU}{z}\right)^t P \tag{34}$$

valid for $|z| > \|QUQ\|$. Since $\psi(r) = 0$, we have $PU(QU)^k P = 0$ for every $k > 0$, and so

$$(QU)^k QUP = U^k QUP. \tag{35}$$

Hence

$$C(z) = \frac{1}{z} \sum_{t=0}^{\infty} \left(\frac{U}{z}\right)^t QUP. \tag{36}$$

Using the results of section 2.2, we can show that

$$U^t QUP\rho(r, \omega) = \frac{1}{N} \sum_{p=1}^N e^{i\theta_p r} e^{i\theta_p(s_{(1-2t)}+s_{(3-2t)}+\dots+s_{-1})} (e^{i\theta_p s_{(-1-2t)}} - \cos \theta_p) \langle \rho_\theta(\omega) \rangle. \tag{37}$$

Combining equations (36) and (37), we obtain

$$C(z)\rho(r, \omega) = \lim_{K \rightarrow \infty} \frac{1}{N} \sum_{p=1}^N e^{i\theta_p r} \langle \rho_\theta(\omega) \rangle A_K(z) \tag{38}$$

where

$$\begin{aligned} A_K(z) &= \frac{1}{z} \sum_{t=0}^K e^{i\theta_p(s_{(1-2t)}+s_{(3-2t)}+\dots+s_{-1})} (e^{i\theta_p s_{(-1-2t)}} - \cos \theta_p) z^{-t} \\ &= \frac{1}{z} (z - \cos \theta_p) \sum_{t=0}^{K-1} \frac{1}{z^t} e^{i\theta_p(s_{-1}+s_{-3}+\dots+s_{(1-2t)})} + \frac{1}{z^{K+1}} e^{i\theta_p(s_{-1}+s_{-3}+\dots+s_{(1-2K)})} - 1. \end{aligned} \tag{39}$$

Hence

$$\begin{aligned} [P + C(z)]\rho(r, \omega) &= \lim_{K \rightarrow \infty} \frac{1}{N} \sum_{p=1}^N e^{i\theta_p r} \langle \rho_\theta(\omega) \rangle \left[z^{-K} e^{i\theta_p(s_{-1}+s_{-3}+\dots+s_{(1-2K)})} \right. \\ &\quad \left. + \sum_{t=0}^{K-1} \frac{1}{z^{t+1}} (z - \cos \theta_p) e^{i\theta_p(s_{-1}+s_{-3}+\dots+s_{(1-2t)})} \right]. \end{aligned} \tag{40}$$

A similar calculation shows that

$$[D(z) + P]\rho(r, \omega) = \lim_{K \rightarrow \infty} \frac{1}{N} \sum_{p=1}^N e^{i\theta_p r} \left[z^{-K} \langle e^{i\theta_p(s_{-1} + s_{-3} + \dots + s_{(1-2K)})} \rho_{\theta_p}(B^{-2K}\omega) \rangle \right. \\ \left. + \sum_{t=0}^{K-1} \frac{(z - \cos \theta_p)}{z^{t+1}} \langle e^{i\theta_p(s_{-1} + s_{-3} + \dots + s_{(1-2t)})} \rho_{\theta_p}(B^{-2t}\omega) \rangle \right]. \quad (41)$$

To be meaningful, the power series above must converge strongly in the Banach space of bounded linear operators from the space of ensembles into itself. The resulting expressions will be evaluated at the poles of $W(z)$. As we can see from equation (34), the series converge when $|z| > \|QUQ\|$. So if the poles of $W(z)$ lie in this region, the expressions derived in this section can be used. This is the same condition as the one mentioned at the end of the previous section.

The above equations, although complex, lead to a relatively simple form for the asymptotic operators which are the aim of the theory.

3.5. The asymptotic evolution operator

We are now ready to follow the procedure described in section 3.3, and construct the asymptotic operators $\Sigma(z)$ and Σ_t . From equation (21), we can see that the residues W_j of the P -subspace resolvent are operators which act on ensembles ρ as follows

$$W_p \rho(\omega, r) = \frac{1}{N} e^{ir\theta_p} \langle \rho_{\theta_p} \rangle. \quad (42)$$

There is one residue for each $z_p = \cos \theta_p$ with $|z_p| > \lambda$.

We now use equation (31) to construct $\Sigma(z)$. The orthogonality of Fourier components leads to a great simplification, and we find that $\Sigma(z)$ can be written as

$$\Sigma(z)\rho(\omega, r) = \frac{1}{N} \sum_{|\cos \theta_p| > \lambda} e^{ir\theta_p} h_{\theta_p}(\omega) \frac{1}{z - \cos \theta_p} \langle g_{\theta_p} \rho_{\theta_p} \rangle \quad (43)$$

where the sum is over all values of p such that $|\cos \theta_p| > \lambda$, and g_θ and h_θ are defined by

$$g_\theta(\omega) = \lim_{m \rightarrow \infty} (\cos \theta)^{-m} e^{i\theta(s_1 + s_3 + \dots + s_{(2m-1)})}. \quad (44)$$

and

$$h_\theta(\omega) = \lim_{m \rightarrow \infty} (\cos \theta)^{-m} e^{i\theta(s_{-1} + s_{-3} + \dots + s_{(1-2m)})}. \quad (45)$$

The inverse transform (equation (32)) gives us the time-dependent operator

$$\Sigma_t \rho(\omega, r) = \frac{1}{N} \sum_{|\cos \theta_p| > \lambda} e^{ir\theta_p} h_{\theta_p}(\omega) \cos^t \theta_p \langle g_{\theta_p} \rho_{\theta_p} \rangle. \quad (46)$$

Thus for finite N , the long-time behaviour of ensembles is described by a sum of independent, exponentially decaying modes.

The meaning of the limits in equations (44) and (45) will be clarified in section 4. The distribution h_θ is an element of the space S_μ^θ , and g_θ is an element of the dual to S_μ^θ ; that is, a linear functional acting on this space. The limits are defined in the topologies of these spaces.

4. The space of allowed ensembles

In this section, we define the set of ensembles $\rho(\omega, r)$ for which the subdynamics theory can be rigorously justified. The set is restricted by a smoothness condition, but as well as functions, it contains elements which are not functions, but must be considered as distributions (linear functionals on a dual space). The definition of this function space will allow us to give answers to two questions.

The first is a technical question. In section 3, we made two assumptions which have not yet been justified. The first assumption was that the evolution operator U is bounded, so that the power-series expansion (20) of the resolvent is correct for large enough z . The second assumption was that QUQ is bounded, and that its bound is small enough so that some singularities of $PR(z)P$ lie in the region $|z| > \|QUQ\|$. In this section, we shall show that these conditions are satisfied when the operators act on ensembles in a certain space.

The second question is concerned with physical interpretation. The DRW is deterministic and reversible in the same way as the Hamiltonian systems of classical physics. There is no ‘arrow of time’ in the model. However, we have in Σ_t an exact asymptotic description of the evolution of ensembles which shows a decay towards equilibrium in the future, but not in the past. There is an apparent contradiction here. However, this is resolved by the fact that the set of ensembles for which the asymptotic description is valid is not time symmetric. The distinction between past and future enters the theory in this way.

The approach taken in this section is similar to the one used in an earlier paper (Evans and Coveney 1995) to analyse the ‘pastry-cook’ system. There, we separated ensembles into ‘even’ and ‘odd’ parts ϕ_+ and ϕ_- , which evolve independently. Here we make use of the fact that in the DRW, the Fourier components $\rho_\theta(\omega)$ of an ensemble evolve independently. We therefore consider each Fourier component separately in our definition of the space of allowed ensembles.

One basic idea underlies the mathematics of this section. Since the shifting of the variable ω is central to the dynamics of our model system, we should find exponential decay of correlations with time if the dependence of ensemble functions on the elements ω_i decays exponentially with $|i|$. When ω is translated into coordinates x and y , such functions become Lipschitz continuous in the (x, y) plane. The definition of the functional K_μ^θ in equation (51) is motivated by this idea. It will therefore come as no surprise when, in section 4.4, we find that Lipschitz continuous functions are part of the function space.

4.1. Definition of the function space

The Fourier transform of an ensemble $\rho(\omega, r)$ is defined by equation (7). Since Fourier components evolve independently, we can define an operator U_θ for each value of θ , which carries a Fourier component forward in time by one unit. If the Fourier component of an ensemble for angle θ at time t is $\rho_\theta(\omega)$, then the same component at time $t + 1$ is $U_\theta \rho_\theta(\omega)$. The family of operators U_θ is defined by

$$U_\theta f(\omega) = e^{is_{-1}(\omega)\theta} f(B^{-2}\omega) \tag{47}$$

where $f(\omega)$ is a function of position in the unit square.

The action of the projection operators P and Q on a Fourier component ρ_θ is independent of θ :

$$Pf(\omega) = \langle f \rangle \tag{48}$$

where $\langle f \rangle$ denotes the constant function multiplied by the mean of f over the unit square, and

$$Qf(\omega) = f(\omega) - \langle f \rangle. \quad (49)$$

We now define a functional acting on functions $f(\omega)$, which will be used to create a norm for our vector space of ensembles. To avoid cluttered equations, we use the notation

$$\sup_{\omega, \delta\omega|n} |\delta[A(\omega)]| \equiv \sup_{\omega, \omega': \omega_i = \omega'_i \ (i \leq n)} |A(\omega) - A(\omega')|. \quad (50)$$

In this equation, the supremum is over all pairs of bi-infinite strings of symbols ω, ω' which agree on symbols ω_i with $i \leq 0$.

Let

$$K_\mu^\theta[f] = \sup_{k \in \mathcal{Z}} \mu^{-k} \delta_k^\theta[f] \quad (51)$$

where \mathcal{Z} denotes the set of integers, μ is a real number in the range $0 < \mu < 1$ and

$$\delta_k^\theta[f] = \sup_{\omega, \delta\omega|0} |\delta[U_\theta^k f(\omega)]|. \quad (52)$$

The functional $\delta_k^\theta[f]$ states how much the function $U_\theta^k f(\omega)$ depends on symbols ω_i to the right of the zeroth in ω .

Another way of writing this equation is, for $k < 0$,

$$\delta_k^\theta[f] = \sup_{\omega, \delta\omega|2k} |\delta[e^{-i\theta(s_{-1}(\omega) + s_{-3}(\omega) + \dots + s_{(1+2k)}(\omega))} f(\omega)]|. \quad (53)$$

For $k \geq 0$, we have

$$\delta_k^\theta[f] = \sup_{\omega, \delta\omega|2k} |\delta[e^{i\theta(s_1(\omega) + s_3(\omega) + \dots + s_{(2k-1)}(\omega))} f(\omega)]|. \quad (54)$$

In the latter case, since the difference $\delta\omega$ involves only symbols ω_i to the right of $i = k$, this simplifies to

$$\delta_k^\theta[f] = \sup_{\omega, \delta\omega|2k} |\delta f(\omega)|. \quad (55)$$

As we saw in section 2.1, the variable $\omega = \dots \omega_{-1} \omega_0 \omega_1 \dots$ is the binary expansion of the y and x coordinates, with $\omega_1 \omega_2 \omega_3 \dots$ being the expansion of x . For $k > 0$ the functional $\delta_k^\theta[f]$ measures how much the function f depends on the digits to the right of the $2k$ th in this expansion of x . For a smooth function, we expect this dependence to decay towards zero as k becomes large. The functional $K_\mu^\theta[f]$ is only finite if this decay is faster than μ^k . Therefore the condition

$$K_\mu^\theta[f] < \infty \quad (56)$$

is a smoothness condition on the x -dependence of $f(\omega)$.

The crucial property of K_μ^θ , which makes it possible to justify the construction of the subdynamics operators, is contained in the following equation:

$$K_\mu^\theta[U_\theta f] = \mu K_\mu^\theta[f]. \quad (57)$$

This is a consequence of the equation

$$\delta_k^\theta[U_\theta f] = \delta_{k+1}^\theta[f] \quad (58)$$

which follows from the definition of δ_k^θ (55). We also note that the functional K_μ^θ is sublinear. That is,

$$K_\mu^\theta[a + b] \leq K_\mu^\theta[a] + K_\mu^\theta[b]. \quad (59)$$

We now define a vector space T_μ^θ , whose completion S_μ^θ will form part of the Banach space of ensemble densities used in the theory.

$$T_\mu^\theta = \left\{ f(\omega) : \begin{array}{l} K_\mu^\theta[Qf] < \infty, |\langle f \rangle| < \infty \\ \text{and } f(\omega) \text{ independent of } \omega_i \text{ for } i < p(f) \end{array} \right\}. \quad (60)$$

The set T_μ^θ contains functions which satisfy the smoothness condition (56), which have a finite mean, and which are independent of all digits ω_i to the left of some point p in the sequence ω . The point p is not fixed, but may depend on the function f . At this stage, μ is still a free parameter, whose value we can choose in the range $0 < \mu < 1$.

The functional

$$\|f\|_\theta = |\langle f \rangle| + K_\mu^\theta[Qf] \quad (61)$$

is a norm on the space T_μ^θ . The proof of this fact is very similar to the proof of the theorem in appendix 1 of Evans and Coveney (1995), so we omit it here. The normed space thus defined is not complete, so we define the space S_μ^θ to be the completion of T_μ^θ under the norm $\|\cdot\|_\theta$ (that is, the space of equivalence classes of Cauchy sequences in T_μ^θ).

The space of ensembles for the DRW is defined in terms of the family S_μ^θ of Banach spaces. For a fixed value of N , an ensemble $\rho(\omega, r)$ is in the space S_μ^N of allowed ensembles when, for every integer j between 1 and N ,

$$\rho_{\theta_j}(\omega) \in S_\mu^{\theta_j} \quad (62)$$

whereas in section 2.2, $\theta_j = 2\pi j/N$. In other words, S_μ^N is the space of ensembles $\rho(\omega, r)$ for which each Fourier component is in the appropriate space. The norm for this space is simply the mean of the norms of all the Fourier components:

$$\|\rho(\omega, r)\| = \frac{1}{N} \sum_{j=1}^N \|\rho_{\theta_j}\|_{\theta_j}. \quad (63)$$

4.2. Bounds on U and QUQ

In this section, we derive bound on the operators U and QUQ . These two bounds are the conditions needed to justify the steps in the application of the subdynamics formalism in section 3.

The bound on U is also necessary for another reason. If an operator is bounded and linear on a normed space such as T_μ^θ , it can be extended uniquely to the completed space S_μ^θ by the so-called BLT theorem (see, for example, Reed and Simon (1972), theorem I.7). It is clear from the definition of the norm that P and Q are bounded, and can therefore be extended in this way. Since the functional K_μ^θ is sublinear and continuous with respect to the norm $\|\cdot\|_\theta$, it can be similarly extended to a continuous functional which satisfies equation (57) for any f in the completed space. We may therefore refer to the extension of these operators to S_μ^θ by the same symbols as the original operators, without the risk of confusion. However, it is necessary to prove that U is bounded before the same can be said of the extension of U .

We first give two lemmas which will be useful in deriving the bounds on U and QUQ . Both lemmas are proved in the appendix.

Lemma 4.1. For any constant function A ,

$$K_\mu^\theta[A] \leq 2\mu|A|. \quad (64)$$

Lemma 4.2. For any function $g(\omega)$ for which $\langle g \rangle = 0$,

$$|\langle U_\theta g \rangle| \leq \frac{1}{2} |\sin \theta| K_\mu^\theta [g]. \tag{65}$$

We now prove that U_θ is bounded.

Theorem 4.1. For any f in the space T_μ^θ ,

$$\|U_\theta f\| \leq 5\|f\|. \tag{66}$$

To prove this result, we note that according to the definition of the norm, equation (61),

$$\|U_\theta f\| = |\langle U_\theta f \rangle| + K_\mu^\theta [QU_\theta f]. \tag{67}$$

We first deal with the second term.

$$\begin{aligned} K_\mu^\theta [QU_\theta f] &\leq K_\mu^\theta [U_\theta f] + K_\mu^\theta [PU_\theta f] \\ &= \mu K_\mu^\theta [f] + K_\mu^\theta [\langle U_\theta f \rangle] \\ &\leq \mu K_\mu^\theta [Qf] + \mu K_\mu^\theta [Pf] + K_\mu^\theta [\langle U_\theta f \rangle] \\ &\leq \mu K_\mu^\theta [Qf] + 2\mu^2 |\langle f \rangle| + 2\mu |\langle U_\theta f \rangle| \end{aligned} \tag{68}$$

where the last inequality follows from lemma 4.1. The norm of $U_\theta f$ therefore has the following bound:

$$\|U_\theta f\| \leq |\langle U_\theta f \rangle| (1 + 2\mu) + 2\mu^2 |\langle f \rangle| + \mu K_\mu^\theta [Qf]. \tag{69}$$

A bound on the first term on the right-hand side of this inequality is given by

$$\begin{aligned} |\langle U_\theta f \rangle| &\leq |\langle U_\theta Pf \rangle| + |\langle U_\theta Qf \rangle| \\ &\leq |\langle f \rangle| + \frac{1}{2} |\sin \theta| K_\mu^\theta [Qf] \end{aligned} \tag{70}$$

where the first term follows from the definition of U_θ and the second from lemma 4.2. We therefore have

$$\begin{aligned} \|U_\theta f\| &\leq (1 + 2\mu + 2\mu^2) |\langle f \rangle| + (\mu + \frac{1}{2}(1 + 2\mu) |\sin \theta|) K_\mu^\theta [Qf] \\ &\leq 5\|f\| \end{aligned} \tag{71}$$

since both μ and $\sin \theta$ have magnitude less than 1. Hence U_θ is bounded.

Theorem 4.2. For any $f(\omega)$ in the space S_μ^θ ,

$$\|QU_\theta Qf\| \leq 2\mu \|f\|. \tag{72}$$

From the definition of the norm (61), and equation (57), we have

$$\begin{aligned} \|QU_\theta Qf\|_\theta &= K_\mu^\theta [QU_\theta Qf] \\ &\leq K_\mu^\theta [U_\theta Qf] + K_\mu^\theta [PU_\theta Qf] \\ &\leq \mu K_\mu^\theta [Qf] + 2\mu |\langle U_\theta Qf \rangle| \end{aligned} \tag{73}$$

where the last equality holds because for any function g , Pg is a constant function to which lemma 4.1 applies.

We now use lemma 4.2, which applies to the second term in equation (73) since $\langle Qf \rangle = 0$. We obtain

$$\begin{aligned} \|QU_\theta Qf\|_\theta &\leq \mu(1 + \sin \theta) K_\mu^\theta [Qf] \\ &\leq 2\mu K_\mu^\theta [Qf] \\ &\leq 2\mu \|f\|_\theta \end{aligned} \tag{74}$$

and this completes the proof.

Since the norm for ensembles $\rho(\omega, r)$ is defined simply as a sum of norms of Fourier components (equation (63)), the corresponding properties of the operators U and QUQ in the space S_μ follow immediately:

$$\|U\| < 5 \tag{75}$$

and

$$\|QUQ\rho\| \leq 2\mu\|\rho\|. \tag{76}$$

Hence, in terms of the natural operator norm derived from $\|\cdot\|$, we have

$$\|QUQ\| \leq 2\mu. \tag{77}$$

This is the most important property of the function space which we have defined. As we show in the next section, it implies the convergence of the subdynamics approximation to the true behaviour at long times. It also implies the convergence of the power series used to construct $C(z)$ and $D(z)$ in section 3.4.

4.3. Convergence of the asymptotic operator

Since the operators $C(z)$, $D(z)$ and $S(z)$ share the denominator $z - QUQ$, equation (77) implies that all these operators are holomorphic in the region $|z| > 2\mu$. So by choosing the parameter μ so that $2\mu < \lambda$, we can guarantee that $C(z)$, $D(z)$ and $S(z)$ are all holomorphic for $|z| \geq \lambda$. This both justifies the extension of $W(z)$ to obtain $\Sigma(z)$, and shows that the resulting $\Sigma(z)$ gives a correct description of the long-time behaviour, as we now demonstrate.

From equation (29), we can see that the fact that $C(z)$, $D(z)$ and $S(z)$ are holomorphic for $|z| > 2\mu$ also implies that the only singularities in $R(z)$ apart from those in $\Sigma(z)$ are in the disk $|z| \leq 2\mu$. We can therefore write

$$R(z) = \Sigma(z) + \hat{\Sigma}(z) \tag{78}$$

where $\hat{\Sigma}(z)$ is holomorphic in the region $|z| > 2\mu$. Inverting the z -transform, as in equation (32), then gives

$$U^t = \Sigma_t + \hat{\Sigma}_t \tag{79}$$

where $\hat{\Sigma}_t$ is exponentially bounded as an operator on S_μ^N :

$$\|\hat{\Sigma}_t\| \leq B(2\mu)^t \tag{80}$$

where B is a constant. We can therefore write equation (33) in a stronger form which shows how rapidly the long-time behaviour of ensembles converges to the approximation given by Σ_t . This provides the central theorem of the paper.

Theorem 4.3. For any element ρ in the space S_μ^N , the subdynamics approximation $\Sigma_t\rho$ to the long-time behaviour of ρ converges exponentially to the true behaviour: that is,

$$\|U^t\rho - \Sigma_t\rho\| \leq C(2\mu)^t \tag{81}$$

where C is a constant which depends upon the function ρ .

4.4. Functions contained in the space of allowed ensembles

We have established the validity of the subdynamics analysis for the set of ensembles S_μ^N , but it is not yet whether this set contains enough ensembles to be useful. In this section, we describe a class of functions which is a subset of S_μ^N . These are the functions which obey a certain smoothness condition on their x and y dependence in the unit square.

Theorem 4.4. Suppose that m is a non-negative integer, and that, when written as a function of x and y , $f(\omega)$ is Lipschitz continuous with exponent α on each of the squares of side 2^{-m} into which the unit square can be divided; that is, there is a constant K such that

$$|f(x, y) - f(x', y')| \leq K(|x - x'| + |y - y'|)^\alpha \quad (82)$$

when the first m binary digits of x and x' and of y and y' are the same. Then if

$$\mu 2^{2\alpha} > 1 \quad (83)$$

f is in S_μ^θ for every θ in the range $0 \leq \theta < 2\pi$.

We omit the proof of this theorem, which is almost identical to the proof of theorem 5.4 of Evans and Coveney (1995). It is easy to see that, since the space S_μ^N is defined in terms of the function spaces S_μ^θ , functions $\rho(\omega, r)$ which obey the smoothness condition of the above theorem for every value of r are in S_μ^N . This provides a large class of ensemble functions which is a subset of S_μ^N , and hence for which theorem 4.3 is valid.

To make the theory which we have constructed work, we must choose $2\mu < \lambda$, where λ is the cut-off radius for the singularities of $\Sigma(z)$, introduced in section 3.1. From equation (83), we then have

$$\alpha > \frac{1}{2} \left(1 - \frac{\log \lambda}{\log 2} \right). \quad (84)$$

For example, if we choose $\lambda = 1/\sqrt{2}$, then $\alpha > \frac{3}{4}$. It is important that the value of α is not greater than 1, because for $\alpha > 1$ the condition of Lipschitz continuity is very restrictive: it allows only piecewise constant functions. Rearranging equation (84) yields

$$\lambda > 2^{1-2\alpha}. \quad (85)$$

Hence for the theory to apply to functions which are not piecewise constant, we need

$$\lambda > \frac{1}{2}. \quad (86)$$

4.5. When does the theory fail?

Are there any circumstances where the theory fails for the DRW? In an earlier paper (Evans and Coveney 1995), we saw that for a different model system, the pastry-cook's transformation, the theory fails when $q \leq 3$, where q is a parameter describing the model. In the DRW, the only parameter is N , the number of cells.

The theory will fail unless we can choose a value of λ which includes at least one decaying mode in the operator $\Sigma(z)$. We therefore need to find exponentially decaying modes which have eigenvalues z_n satisfying

$$|z_n| > \frac{1}{2}. \quad (87)$$

The eigenvalues of the exponentially decaying modes are the singularities of the operator $PR(z)P$. From equation (21), we can see that these are

$$z_n = \cos \frac{2\pi n}{N} \quad (88)$$

for $n = 1, 2, \dots, N$, where N is the number of cells. For large values of N , there are eigenvalues distributed throughout the interval $-1 < z_n < 1$, including many satisfying (87).

For smaller values of N , there are fewer eigenvalues. In fact, for $N = 1, 2, 3, 4$ and $N = 6$, there are no eigenvalues satisfying (87). In these cases the theory fails. The root of this problem is the fact that for these small values of N , diffusive behaviour does not dominate at long times. Our choice of projection operator P is therefore incorrect for these cases, and hence the theory cannot work. A different choice of P , including some information about the distribution of an ensemble in the (x, y) coordinates, might give correct results. For all other values of $N > 1$, the theory works as it stands.

If we had used only one baker transformation, rather than two successive transformations, when the DRW was defined in section 2.1, the result would have been rather different. Because non-uniformities in the (x, y) distribution decay more slowly in this case, it is not possible to isolate diffusive modes which dominate the long-time behaviour, for any value of N . In place of equation (86), we obtain $\lambda > 1$, so that there is no way to find an exponentially decaying mode with an eigenvalue $z_i > \lambda$. When only one Baker transformation is used, the theory works only when the set of allowed ensembles is so restrictive that it includes only piecewise constant functions.

5. The thermodynamic limit

So far, we have been working with the version of the DRW which has a finite value of N , so that correlation functions decay exponentially with time. Using the subdynamics formalism, we have constructed operators which give an exact description of the long-time behaviour of ensembles as a superposition of a finite set of exponentially decaying modes. However, our results are expressed in a form which makes taking the limit $N \rightarrow \infty$ simple. In this limit, the inverse Fourier transform, which is a sum for finite N (equation (8)) becomes an integral (equation (9)). This is essentially the only change. Thus equation (46), which gives the action of the asymptotic operator Σ_t for finite N , becomes in the limit $N \rightarrow \infty$

$$\Sigma_t \rho(\omega, r) = \int_{|\cos \theta| > \lambda} d\theta e^{ir\theta} h_\theta(\omega) \cos^t \theta \langle g_\theta \rho_\theta \rangle. \tag{89}$$

As the sum over a finite number of exponentially decaying modes becomes an integral over an infinite number of modes, the long-time behaviour of correlation functions no longer needs to be exponential. It is clear that correlations of pure Fourier components will decay exponentially, but these are exceptional (and represent non-normalizable ensembles).

5.1. Two examples of non-exponential decay

We define a correlation function as the matrix element

$$g_{\phi\psi}(t) = \sum_r \langle \phi(\omega, r) U^t \psi(\omega, r) \rangle. \tag{90}$$

The physical interpretation is that if we prepare an ensemble $\psi(\omega, r)$ at time zero, then the expectation value of the phase-space variable $\phi(\omega, r)$ at time t is $g_{\phi\psi}(t)$.

Equation (89) gives an expression for the long-time limit $\Sigma_{\phi\psi}(t)$ of such correlation functions:

$$g_{\phi\psi}(t) \xrightarrow{t \rightarrow \infty} \Sigma_{\phi\psi}(t)$$

$$= \frac{1}{2\pi} \int_{|\cos \theta| > \lambda} d\theta \cos^t \theta \langle \phi_\theta^* h_\theta \rangle \langle g_\theta \psi_\theta \rangle \quad (91)$$

where $\phi_\theta(\omega)$ and $\psi_\theta(\omega)$ are the Fourier transforms of the functions $\phi(\omega, r)$ and $\psi(\omega, r)$ (equation (7)).

The simplest example of non-exponential decay is the autocorrelation $g_{\delta\delta}(t)$ of the delta function

$$\delta(r) = \begin{cases} 1 & \text{if } r = 0 \\ 0 & \text{otherwise.} \end{cases} \quad (92)$$

The Fourier transform of this function is just a constant:

$$\delta_\theta(\omega) = 1. \quad (93)$$

Evaluating the integrals represented by angled brackets $\langle \dots \rangle$ in equation (91) is simple. Using the definition of g_θ (equation (45)), we obtain

$$\begin{aligned} \langle g_\theta \delta_\theta \rangle &= \langle g_\theta \rangle \\ &= \lim_{m \rightarrow \infty} (\cos \theta)^{-m} \langle e^{i\theta(s_1 + s_3 + \dots + s_{2m-1})} \rangle \\ &= 1 \end{aligned} \quad (94)$$

since $\langle e^{i\theta s_1} \rangle = \cos \theta$. A similar calculation shows that $\langle \delta_\theta^* h_\theta \rangle = 1$. The expression for the long-time limit of correlation functions (91) becomes simply

$$\Sigma_{\delta\delta}(t) = \frac{1}{2\pi} \int_{|\cos \theta| > \lambda} d\theta e^{ir\theta} \cos^t \theta. \quad (95)$$

The limiting behaviour of this integral for large t is easy to obtain using a saddle-point technique. The two saddle points are at $\theta = 0$ and $\theta = \pi$. These points are included in the domain of integration whatever value of λ we choose in the range $0 < \lambda < 1$, so the asymptotic behaviour is independent of λ , as one might expect. The result is

$$g_{\delta\delta}(t) \xrightarrow{t \rightarrow \infty} \frac{1}{(2\pi t)^{1/2}} (1 + (-1)^t). \quad (96)$$

A less trivial example is provided by the autocorrelation of the function

$$\phi(\omega, r) = 3\delta(r)x^2. \quad (97)$$

We can write the x -coordinate in the unit square in terms of ω , since the digits of ω contain the binary expansion of x :

$$x = \sum_{j=1}^{\infty} \omega_j 2^{-j}. \quad (98)$$

Hence, by writing the Fourier transform of ϕ as

$$\phi_\theta(\omega) = 3 \sum_{j,k=1}^{\infty} 2^{-(j+k)} \omega_j \omega_k \quad (99)$$

we can evaluate the two integrals in equation (91):

$$\langle g_\theta \phi_\theta \rangle = \frac{3 + e^{2i\theta}}{(1 + e^{2i\theta})^2} \quad (100)$$

and

$$\langle \phi_\theta^* f_\theta \rangle = 1. \quad (101)$$

The asymptotic expression for the autocorrelation function $g_{\phi\phi}(t)$ is therefore

$$\Sigma_{\phi\phi}(t) = \frac{1}{4\pi} \int_{|\cos\theta| < \lambda} d\theta \cos^t \theta \left(\frac{5 + 6 \cos 4\theta + 5 \cos 2\theta}{3 + \cos 4\theta + 4 \cos 2\theta} \right). \quad (102)$$

By comparing this expression with the autocorrelation of the δ -function (equation (95)), we can see how the distribution over the unit square influences an ensemble's long-time behaviour. For the two functions δ and ϕ which we have considered, the autocorrelation functions behave differently. However, using the saddle-point technique to evaluate the integral in equation (102) one can show that in the long-time limit, $\Sigma_{\phi\phi}(t)$ converges to the same asymptotic form as $\Sigma_{\delta\delta}(t)$ (equation (96)).

6. Conclusion

Since each individual mode in subdynamics decays exponentially, it is not obvious how the theory can be used to analyse systems where correlations decay non-exponentially with time. In this paper we have shown how non-exponential decay arises from the superposition of a continuous range of exponentially decaying modes.

The treatment of the 'thermodynamic' limit $N \rightarrow \infty$ given here has not been mathematically rigorous. This is an example of a general problem. As far as we are aware all attempts to treat systems where the evolution operator has a continuous spectrum using the subdynamics formalism have, as here, taken the approach of starting with a system having a discrete spectrum, and then assuming that the results remain valid in the continuum limit (Petrosky and Hasegawa 1989). This is also true of many other treatments of non-exponential decay in non-equilibrium statistical physics (e.g. Reichl (1980)).

In the case of finite N , we have derived a consistent and exact description of the long-time evolution of ensembles, which includes the irreversible approach to equilibrium. This description is valid for a certain space of non-equilibrium ensembles, and the norm of this space provides the definition of convergence at long times. The choice of this set of ensembles is a crucial part of the theory. The model itself is invertible and time symmetric, and the time asymmetry which ensures an approach to equilibrium is supplied by the ensembles.

The general principle behind the Brussels formalism is that the 'kinetic' properties of a system, which are the macroscopic features accessible to observation, emerge as the dominant long-time behaviour of an exact description. Thus, the kinetic equation that we have derived is not an approximation, but a limiting description to which the system converges when left to itself for long times. If this approach could be carried out with mathematical rigour for more realistic systems, it would offer an attractive alternative to the work of Lanford (1975) and Spohn (1980, 1991), where kinetic equations are derived in a limiting case and are valid only for short times.

One lesson that we have learnt from our studies of discrete-time dynamics, and which can be applied to more complex systems, is that it is easy to be led astray by the mathematical apparatus of the Brussels formalism. The formalism can easily produce incorrect results if the definition of the reduced description does not reflect the long-time dynamics of the system. A mathematical framework for the theory ensures the correctness of the results if it is introduced carefully, but it is important to make sure that the framework includes the functions which we wish to use as ensembles. Otherwise, as in the cases of the pastry-cook's transformation for low values of q (Evans and Coveney 1995), and the DRW for low values of N , we produce a mathematically correct theory which is physically empty.

The system which we have analysed provides a simple illustration of the way macroscopic diffusion arises from chaotic processes at the microscopic level. This phenomenon is ubiquitous in physics, and we hope that our results give an insight into the statistical mechanics of diffusion processes in general.

Acknowledgments

We wish to thank Oscar Bandtlow and Jonathan Wattis for many useful comments. Part of the work described here was supported by the UK Engineering and Physical Sciences Research Council.

Appendix. Proof of lemmas 4.1 and 4.2

Proof of lemma 4.1. For $k \geq 0$, since $\delta_k^\theta = \delta_k^+$, we have

$$\delta_k^- [A] = 0. \quad (103)$$

But for $k < 0$,

$$\delta_k^- [A] \leq 2|A|. \quad (104)$$

Hence

$$K_\mu^- [A] \leq \sup_{k < 0} \mu^{-k} 2|A| = 2\mu|A|. \quad (105)$$

This completes the proof. \square

Proof of lemma 4.2. Let

$$a = \langle (1 - \omega_1)g \rangle \quad (106)$$

and

$$b = \langle \omega_1 g \rangle. \quad (107)$$

Since $\langle g \rangle = 0$,

$$a + b = 0. \quad (108)$$

Now let ω , $\omega + \delta\omega$ be sequences which differ only in the value of bit number 1. By the definition of K_μ^θ (equation (51)),

$$|g(\omega + \delta\omega) - g(\omega)| \leq K_\mu^\theta [g]. \quad (109)$$

Multiplying by ω_1 and integrating, we find

$$\langle \omega_1 |g(\omega + \delta\omega) - g(\omega)| \rangle \leq \frac{1}{2} K_\mu^\theta [g] \quad (110)$$

and hence

$$|\langle \omega_1 g(\omega + \delta\omega) \rangle - \langle \omega_1 g(\omega) \rangle| \leq \frac{1}{2} K_\mu^\theta [g]. \quad (111)$$

But

$$\langle \omega_1 g(\omega + \delta\omega) \rangle = \langle (1 - \omega_1)g(\omega) \rangle \quad (112)$$

so we have

$$|b - a| \leq \frac{1}{2} K_\mu^\theta [g]. \quad (113)$$

Finally,

$$\begin{aligned}\langle U_\theta g \rangle &= \langle e^{is(\omega_1)\theta} g \rangle \\ &= e^{i\theta} a + e^{-i\theta} b \\ &= (a + b) \cos \theta + i(a - b) \sin \theta.\end{aligned}\tag{114}$$

So, using equations (108) and (113), we obtain

$$|\langle U_\theta g \rangle| \leq \frac{1}{2} K_\mu^\theta [g] |\sin \theta|.\tag{115}$$

This completes the proof. \square

References

- Antoniou I E and Tasaki S 1993 Generalized spectral decompositions of mixing dynamical systems. *Int. J. Quantum Chem.* **46** 425–74
- Arnold V I and Avez A 1968 *Ergodic Problems of Classical Mechanics* (New York: Benjamin)
- Balescu R 1975 *Equilibrium and Non-equilibrium Statistical Mechanics* (New York: Wiley)
- Bandtlow O F and Coveney P V 1994 On the discrete time version of the Brussels formalism *J. Phys. A: Math. Gen.* **27** 7939–55
- Coveney P V and Evans A K 1994 Canonical non-equilibrium ensembles and subdynamics *J. Stat. Phys.* **77** 229–58
- Elskens Y and Kapral R 1985 Reversible dynamics and the macroscopic rate law for a solvable Kolmogorov system: the three bakers' reaction *J. Stat. Phys.* **38** 1027–49
- Evans A K and Coveney P V 1995 On exponential long-time evolution in statistical mechanics *Proc. R. Soc. A* **448** 293–319
- Garrido P L and Gallavotti G 1994 Billiards correlation functions *J. Stat. Phys.* **76** 549–85
- Gaspard P 1992 Diffusion, effusion and chaotic scattering: an exactly solvable Liouvillian dynamics *J. Stat. Phys.* **68** 673–747
- Gaspard P 1993 What is the role of chaotic scattering in irreversible processes? *Chaos* **3** 427–42
- Giesel F and Nierwetberg J 1982 Onset of diffusion and universal scaling in chaotic systems *Phys. Rev. Lett.* **48** 7–10
- Hasegawa H H and Driebe D J 1994 Intrinsic irreversibility and the validity of the kinetic description of chaotic systems *Phys. Rev. E* **50** 1781–809
- Hasegawa H H and Saphir W C 1992 Unitarity and irreversibility in chaotic systems *Phys. Rev. A* **46** 7401–23
- Lanford O E 1975 *Time Evolution of Large Classical Systems* (Berlin: Springer)
- Obcemea C and Brändas E 1983 Analysis of Prigogine's theory of subdynamics *Ann. Phys.* **151** 383–430
- Penrose O and Coveney P V 1994 Is there a canonical non-equilibrium ensemble? *Proc. R. Soc. A* **447** 631–46
- Petersen K 1983 *Ergodic Theory* (Cambridge: Cambridge University Press)
- Petrosky T Y and Hasegawa H 1989 Subdynamics and nonintegrable systems *Physica* **160A** 351–85
- Reed M and Simon B 1972 *Methods of Modern Mathematical Physics. Vol. I: Functional Analysis* (New York: Academic)
- Reichl L E 1980 *A Modern Course in Statistical Physics* (London: Edward Arnold)
- Schuster H G 1988 *Deterministic Chaos: An Introduction* (Weinheim: Physik)
- Spohn H 1980 Kinetic equations from Hamiltonian dynamics: Markovian limits *Rev. Mod. Phys.* **52** 569–616
- Spohn H 1991 *Large Scale Dynamics of Interacting Particles* (Berlin: Springer)
- Vollmer J, Tel T and Breyman W 1997 Equivalence of irreversible entropy production in driven systems: an elementary chaotic map approach *Phys. Rev. Lett.* **79** 2759–62