Some Examples of Simple Systems*

D. A. Lavis
Department of Mathematics
King's College, Strand, London WC2R 2LS - U.K.
Email:David.Lavis@kcl.ac.uk;
Webpage:www.mth.kcl.ac.uk/∼dlavis

1 Introduction

The aim of this work is to provide a set of examples which may give some insight into the behaviour of 'real' statistical mechanical systems. The choice of examples is motivated partly by reasons of precedence (I have tried to include the examples which occur most frequently in the literature) and partly to illustrate particular aspects of the area of interest. Our main purpose is to provide exact[1] simulations using MAPLE. The investigation of kinetic equations (see e.g. Bogoliubov, 1962) is, of course, important, but they are, in general, approximations to the behaviour of the underlying systems. As such any conclusions which can be drawn are of limit use for the discussion of foundational problems. However, the restriction to exact simulations is severe. Most interesting results in statistical mechanics are for cooperative systems and even at equilibrium there are only a few exact results (see e.g. Lavis and Bell, 1999) and no many-body cooperative systems are tractable in terms of following their exact dynamic behaviour. In addition to the restriction to systems of non-interacting microsystems we are also forced to consider only linear systems. Calculations for the evolution of non-linear equations, like the assembly of anharmonic oscillators discussed by Bricmont (2001), are normally some type of perturbation expansion (see e.g. Drazin, 1992) and so of the same limited use as kinetic equations. Thus we shall consider assemblies of non-interacting microsystems driven by linear equations. We never-the-less hope to demonstrate that these produce interesting behaviour, which serves to clarify some of the contentious issues in statistical mechanics.

For all of our examples we shall investigate the time-evolution of the *Boltzmann entropy* and, in the case of the spin-echo system in Sec. 6, we also consider the fine-grained and course-grained versions of the *Gibbs entropy*.

1.1 The Boltzmann Entropy

Consider a system, which at time t has microstate given by the vector $\boldsymbol{x}(t)$ in the phase-space Γ.[2]

[1] See Sec. 7.

[2] Both time and the space Γ may be continuous or discrete and the 'dynamics' which drives $\boldsymbol{x}(t)$ in Γ will be either deterministic, or could be taken to be stochastic.

Macrostates (observable states) are defined by a set Ξ of macroscopic variables.[3] Let the set of macrostates be $\{\mu\}_{\Xi}$. They are so defined that every $\boldsymbol{x} \in \Gamma$ is in exactly one macrostate denoted by $\mu(\boldsymbol{x})$ and the mapping $\boldsymbol{x} \to \mu(\boldsymbol{x})$ is many-one. Every macrostate μ is associated with its 'volume' $\mathcal{V}(\mu) = \mathcal{V}_\mu$ in Γ.[4] We thus have the map $\boldsymbol{x} \to \mu(\boldsymbol{x}) \to \mathcal{V}_\mu(\boldsymbol{x})$ from Γ to $\mathbb{R}^+$ or $\mathbb{N}$. The *Boltzmann entropy* is defined by

$$S_{\mathrm{B}}(\boldsymbol{x}) = k_{\mathrm{B}} \ln[\mathcal{V}_\mu(\boldsymbol{x})/\mathcal{V}_{\min}], \tag{1}$$

where $\mathcal{V}_{\min}$ is the volume of the macrostate of minimum volume.[5] This is a phase function depending on the choice of macroscopic variables Ξ, so how do we expect that it will behave? The point of view of the *typical system approach* (Lebowitz, 1993; Bricmont, 1995) can be expressed as follows:

> *If the system starts at a phase point associated with low entropy then we expect the entropy to increase.*[6] *As it increases there is the possibility of fluctuations, which could be large. When the entropy gets near to its maximum value then it will still be expected to fluctuate and those fluctuations could be large but we don't expect large fluctuations to occur very often.*

This assertion, of course, runs counter to the strict form of the second law, which does not allow any decreases in entropy of an isolated system.

We shall see in our examples that an important feature of the system, which is determined by the dynamics, is the accessibility of one macrostate from another. By μ' being 'accessible' from μ we shall mean *directly* accessible (without any intermediate macrostates) and μ' is accessible from μ if there exists an $\boldsymbol{x} \in \mu$ which maps with time so that the next macrostate it visits *is* (in the case of deterministic dynamics) or *has non-zero probability of being* (in the case of stochastic dynamics) μ'. To clarify this point we introduce a number of definitions. Given a macrostate μ, the set of macrostates which are accessible from μ is denoted by $\mathcal{S}(\mu)$. For reversible dynamic systems[7] it is normally the case that the phase points $\boldsymbol{x}$ and $J\,\boldsymbol{x}$ belong to the same macrostate. Then $S_{\mathrm{B}}(\boldsymbol{x}) = S_{\mathrm{B}}(J\,\boldsymbol{x})$ and if μ' is accessible from μ then μ is accessible from μ'. Now divide $\mathcal{S}(\mu)$ into two subsets $\mathcal{S}^{(+)}(\mu)$ consisting of those macrostates with volumes greater *or equal*[8] to $\mathcal{V}(\mu)$ and $\mathcal{S}^{(-)}(\mu)$ consisting of those macrostates with volumes less than to $\mathcal{V}(\mu)$. Entropy will increase/decrease in the transition $\mu \to \mu'$ if $\mu' \in \mathcal{S}^{(\pm)}(\mu)$ and the proportions of volume accessible from μ leading to increase/decrease of entropy are

$$\mathsf{v}^{(\pm)}(\mu) = \frac{\mathcal{V}(\mathcal{S}^{(\pm)}(\mu))}{\mathcal{V}(\mathcal{S}(\mu))}, \tag{2}$$

[3]These may include some thermodynamic variables (volume, number of particles etc.) but they can also include other variables, specifying, for example, the number of particles in a set of subvolumes. Ridderbos (2002) denotes these by the collective name of *supra-thermodynamic variables*.

[4]In the case where Γ is continuous the volume of μ will normally be its Lebesque measure; when Γ is discrete the volume will be the number of points in μ.

[5]The (optional) inclusion of this factor is convenient both for setting a zero to the entropy and a scale for the volume.

[6]If it is the macrostate of least volume it will, of course, have to initially increase.

[7]Meaning that, there exists an operator J on the points of Γ such that, if $\boldsymbol{x}' = \phi_t \boldsymbol{x}$ then $J^{-1}\phi_t J\,\boldsymbol{x}' = \boldsymbol{x}$.

[8]So we shall be using 'increase' to mean increase or be equal to and decrease in the strict sense.

where $\mathcal{V}(\mathcal{S}(\mu))$ denotes the total volume of the members of $\mathcal{S}(\mu)$ (and similarly for $\mathcal{S}^{(\pm)}(\mu)$). In some of our examples it is simple to compute $\mathsf{v}^{(\pm)}(\mu)$ as functions of $\mathcal{V}$.

1.2 The Gibbs Entropy

Suppose the system consists of N microsystems.[9] The (fine-grained) Gibbs entropy is given by the functional

$$S_{\mathrm{FGG}}[\rho_N(t)] = -k_{\mathrm{B}} \int_{\Gamma_N} \rho_N(\boldsymbol{x};t) \ln\{\rho_N(\boldsymbol{x};t)\} \mathrm{d}\Gamma_N. \tag{3}$$

of the (fine-grained) probability density function $\rho_N(\boldsymbol{x};t)$ on Γ_N. For a measure-preserving system for which $\rho_N(\boldsymbol{x};t)$ satisfies Liouville's equation $S_{\mathrm{FGG}}[\rho_N(t)]$ remains constant with time, as we shall demonstrate explicitly for the spin-echo system of Sec. 6. The resolution to this problem suggested by Gibbs (1902, p. 148) (see also Ehrenfest and Ehrenfest-Afanassjewa, 1912) is to course-grain the phase–space Γ_N, in the manner in which macrostates are usually obtained in the Boltzmann approach. We first note that for a system of non-interacting microsystems the probability density function factorizes into a product of single-microsystem densities. We shall for simplicity, and because it is true for all our examples, suppose that the microsystems are identical. Then

$$\rho_N(\boldsymbol{x};t) = \prod_{i=1}^{N} \rho_1(\mathring{\boldsymbol{x}}^{(i)};t), \tag{4}$$

where $\mathring{\boldsymbol{x}}^{(i)}$, representing the state of the i-th microsystem, moving in one of N identical copies of the single-microsystem phase space Γ_1. Then

$$S_{\mathrm{FGG}}[\rho_N(t)] = -k_{\mathrm{B}} N \int_{\Gamma_1} \rho_1(\mathring{\boldsymbol{x}};t) \ln\{\rho_1(\mathring{\boldsymbol{x}};t)\} \mathrm{d}\Gamma_1. \tag{5}$$

We now divide Γ_1 into an enumerable set of hypercubic cells with edge-length in the direction of the $\mathring{x}_j$ axis of length $\triangle\mathring{x}_j$. The cells, each of volume $\nu = \prod_j \triangle\mathring{x}_j$, are denoted by γ_k and we define the course-grained probability density by

$$\bar{\rho}_1(k;t) = \int_{\gamma_k} \rho_1(\mathring{\boldsymbol{x}};t)\mathrm{d}\Gamma_1 \tag{6}$$

and the course-grained Gibbs entropy by

$$S_{\mathrm{CGG}}[\bar{\rho}_N(t)] = -k_{\mathrm{B}} N \sum_k \bar{\rho}_1(k;t) \ln\{\bar{\rho}_1(k;t)\} + k_{\mathrm{B}} N \ln(\nu). \tag{7}$$

The presence of the (optional) second term in (7) is required if we demand consistency with the fine-grained entropy in the case where the fine-grained density is uniform (with possibly different values) over each of the cells. Then, from (6), $\bar{\rho}_1(k;t) = \nu\rho_1(\mathring{\boldsymbol{x}}_k;t)$, where $\mathring{\boldsymbol{x}}_k$ is any point in γ_k and substituting into (5) gives (7).[10]

[9] In indication of which we denote the phase space by Γ_N.

[10] Alternatively the final term in (7) could be absorbed if the formula were written in the form of an integral (rather than summation) over the piecewise constant course-grained density.

It is not difficult to show that, in general, if we begin with any fine-grained density and calculate the $S_{\mathrm{FGG}}[\rho_N(t)]$, and then apply course-graining and calculate $S_{\mathrm{CGG}}[\bar{\rho}_N(t)]$,

$$S_{\mathrm{FGG}}[\rho_N(t)] \leq S_{\mathrm{CGG}}[\bar{\rho}_N(t)], \tag{8}$$

with equality only if the fine-grained density is uniform over the cells of the course-graining. Now we can conceive of two possible ways of tracing the evolution of entropy in the Gibbs course-grained picture.

(i) We could begin with some fine-grained density giving entropy $S_{\mathrm{FGG}}[\rho_N(0)]$ at $t = 0$ and watch its evolution as time increases.[11] If at time $t' \geq 0$ we course-grain, then

$$S_{\mathrm{FGG}}[\rho_N(0)] = S_{\mathrm{FGG}}[\rho_N(t')] \leq S_{\mathrm{CGG}}[\rho_N(t')]. \tag{9}$$

However if we course-grain at two instants $0 \leq t' < t''$ it is not necessarily the case that

$$S_{\mathrm{CGG}}[\rho_N(t')] \leq S_{\mathrm{CGG}}[\rho_N(t'')] \tag{10}$$

The course-grained entropy will not necessarily show monotonic increase. However, the graph of the course-grained entropy, will not depend on the instants at which course-graining is applied.

(ii) If, instead of the strategy adopted in (i) we course-grain at t' then follow the evolution of the course-grained density and then *re-course-grain* at the later time t'', (10) will hold. Course-grained entropy will show monotonic increase. However, the graph of entropy against time will be affected by the instances at which course-graining is applied.

We comment further on these observation in relation to the spin-echo system in Sec. 7.

2 Simple Markovian and Simple Random-Walk Models

Suppose that a system has m macrostates and that the macrostate $\mu(n)$ for $n = 1, 2, \ldots, m$, consists of $\mathcal{V}(n) = n$ phase points. In the phase space there will be $\frac{1}{2}m(m+1)$ points. Let a trajectory of the system be a sequence of random jumps between phase points. Fig. 1 shows a simulation giving the scaled macrostate volumes for the system with $m = 100$, starting in $\mu(1)$. It is clear that whereas the initial entropy jump is necessarily positive, the subsequent behaviour shows wildly fluctuating behaviour. The band of frequently visited states have scaled volumes spreading over an approximate range $[0.3, 1.0]$ with frequent fluctuations to lower values.

The problem with this simple Markovian (SM) model arises from the fact that we have allowed accessibility between all pairs of macrostates. Given that, for large n, the majority of the phase space will be occupied by points in

[11] Or decreases, but that's another story.

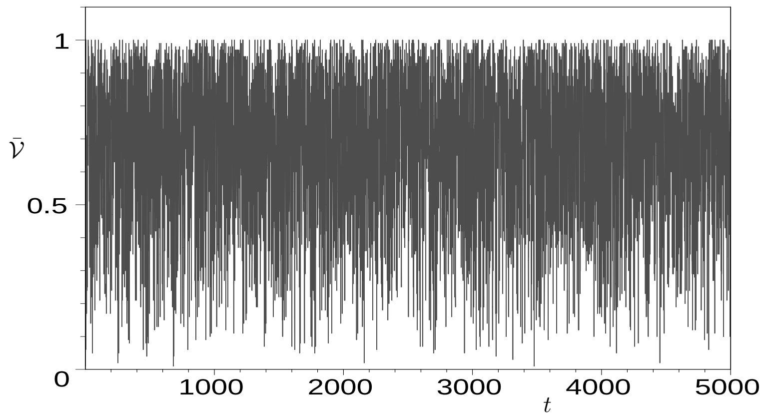


Figure 1: Simulation of the SM model with $m = 100$.

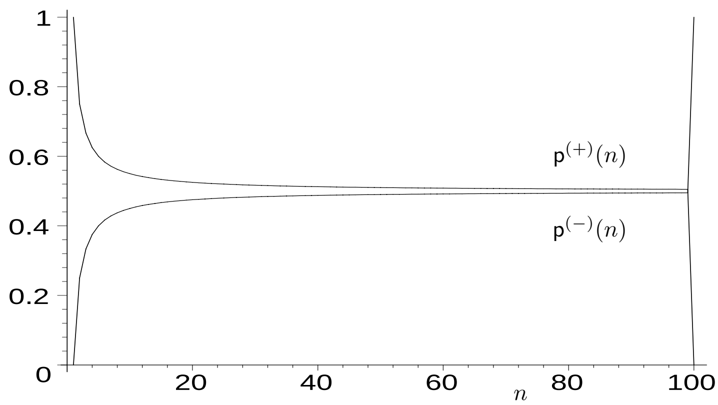


Figure 2: The probabilities $\mathsf{p}^{(\pm)}(n)$, for the SRW model, plotted against n for a the case of $m = 100$ states.

macrostates of smaller volumes,[12] unrestricted transitions will lead to large and persistent fluctuations in macrostate volume and entropy. If on the other hand transitions are allowed from a macrostate n only to macrostates $n \pm 1$ the model becomes a simple random walk (SRW) and the picture changes. The probabilities that, when the system is in macrostate $\mu(n)$, it moves to a macrostate with larger/smaller entropy are

$$\mathsf{p}^{(\pm)}(n) = \pi(n, n \pm 1) = \left\{ \begin{array}{ll} \dfrac{n \pm 1}{2n}, & \text{if } 1 \leq n < m, \\ \frac{1}{2}(1 \mp 1), & \text{if } n = m, \end{array} \right\} = \mathsf{v}^{(\pm)}(n), \qquad (11)$$

[12] For $m = 100$ about half the phase space consists of points in macrostates with $n < 70$.

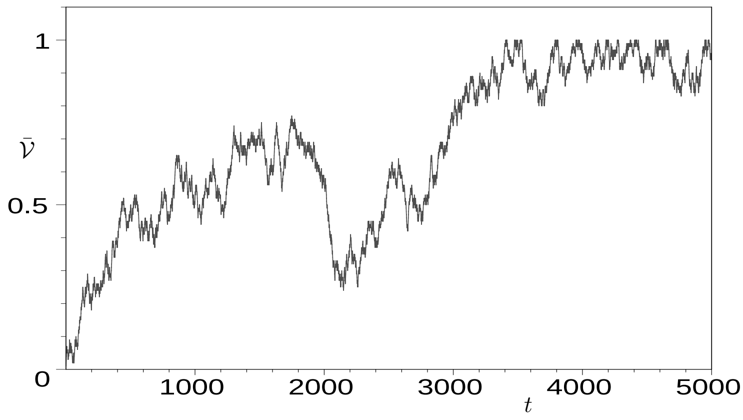


Figure 3: Simulation of the SRW model with $m = 100$.

where[13]

$$\pi(n, n \pm 1) = \mathsf{Prob}[x(t+1) = n \pm 1 | x(t) = n]. \tag{12}$$

Graphs of $\mathsf{p}^{(\pm)}(n)$ are given in Fig. 2. The probability of a transition leading to an increase in entropy falls to a value near to a half, converging with the probability of a drop in entropy, when the system is in a state of volume near to the maximum. Of course, when the system is in the macrostate of maximum volume then any transition must lead to a drop in entropy. These forms for the probabilities illuminate the reduction in fluctuations as macrostate volume (and entropy) increases as shown in Fig. 3.

This behaviour is much closer to our expectations for the entropy as described at the beginning of Sec. 1. Thus we are learning that accessibility between macrostates plays an important part in the behaviour of the model.[14] The probability $\rho(n;t)$ that $x(t) = n$, for $n = 1, 2, \ldots, m$ satisfies[15]

$$\rho(n; t+1) = \rho(n-1; t)\pi(n-1, n) + \rho(n+1; t)\pi(n+1, n), \tag{13}$$

and it is easy to see that

$$\rho^{\star}(n) = \frac{\mathcal{V}(n)}{\displaystyle\sum_{k=1}^{m} \mathcal{V}(k)} = \frac{2n}{m(m+1)} \tag{14}$$

is an equilibrium, meaning time-invariant, solution of (13).

[13] As in Parzen (1962) we use the convention that $\pi(n,k)$ is the transition probability *from* state n *to* state k. A random walk is a Markov chain for which $\pi(n,k) = 0$ if $|k - n| > 1$. In our examples we also have $\pi(n,n) = 0$ for all n.

[14] In the the spin-echo system of Sec. 6 accessibility is determined by the dynamics of the model. A transition between two macrostates occurred whenever the location of one dipole described by the variables (θ, ω) in Γ_1 passed from one cell to another. This, of course, meant that the consequent change in entropy is quite small. Similar contiguity mechanisms also apply to the other models described here.

[15] With $\rho(n;t) = 0$ for all t, when n is not in the range $1 \leq n \leq m$.

3 The Dog-Flea Model

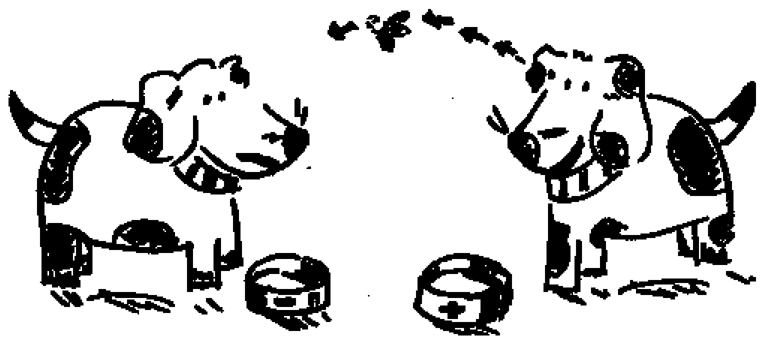
Two of the most well-known 'toy' models of statistical mechanics are the *dog-flea model*[16] of Ehrenfest and Ehrenfest-Afanassjewa (1907) (see also Emch and Liu, 2002, p. 106–112, which contains further references) and the *ring model* of Kac (1959). We shall now see that the dog-flea model is a stochastic version of a special case of the ring model.

Two dogs called Plus and Minus share a population of N fleas, where for convenience we suppose that N is even. Each axis of the phase space Γ_N denotes the state of a flea and has two points $+1$ and -1 indicating that the flea is on Plus or Minus respectively. A macrostate of the system, for fixed N, is specified by one variable n, in the range $\left[-\frac{1}{2}N, \frac{1}{2}N\right]$, where $N^{(+)} = \frac{1}{2}N + n$ and $N^{(-)} = \frac{1}{2}N - n$ are respectively the number of fleas on Plus and Minus and the volume of (number of phase points in) the macrostate is

$$\mathcal{V}(n) = \frac{N!}{\left(\frac{1}{2}N + n\right)!\left(\frac{1}{2}N - n\right)!}. \tag{15}$$

The Boltzmann entropy[17] $S_{\mathrm{B}}(n) = k_{\mathrm{B}} \ln[\mathcal{V}(n)]$ is symmetric in n with maximum at $n = 0$ and zeros at $n = \pm\frac{1}{2}N$.

Now suppose the fleas are numbered 1 to N and the flea-trainer chooses a number at random in this range and orders that flea to change dogs.



This again is a random walk with transition probabilities

$$\pi(n, n \pm 1) = \frac{N \mp 2n}{2N}, \qquad -\tfrac{1}{2}N \le n \le \tfrac{1}{2}N. \tag{16}$$

The probabilities $\mathsf{p}^{(\pm)}(n)$ that entropy and the macrostate volume increase/decrease are the probabilities that $|n|$ decreases/increases respectively. That is

$$\mathsf{p}^{(\pm)}(n) = \begin{cases} \pi(n, n \mp 1), & \text{if } 0 < n \le \frac{1}{2}N, \\ \frac{1}{2}(1 \mp 1), & \text{if } n = 0, \\ \pi(n, n \pm 1), & \text{if } -\frac{1}{2}N \le n < 0, \end{cases} \tag{17}$$

where now the random variable for the system is $x(t) = n$. We denote by $\mathsf{p}^{(+)}(\bar{\mathcal{V}})$ the function $\mathsf{p}^{(+)}(n)$ when, through (15), it is plotted against $\bar{\mathcal{V}}$. These probabilities, together with the scaled volumes given by (2) are shown in Fig. 4. Unlike the SRW model, $\mathsf{v}^{(\pm)}(n) \neq \mathsf{p}^{(\pm)}(n)$. This is because the model has a

[16] A vermin-free version of this model replaces the dogs by urns and the fleas by balls.

[17] In this case, of course, $\mathcal{V}_{\min} = \mathcal{V}\left(\pm\frac{1}{2}N\right) = 1$.

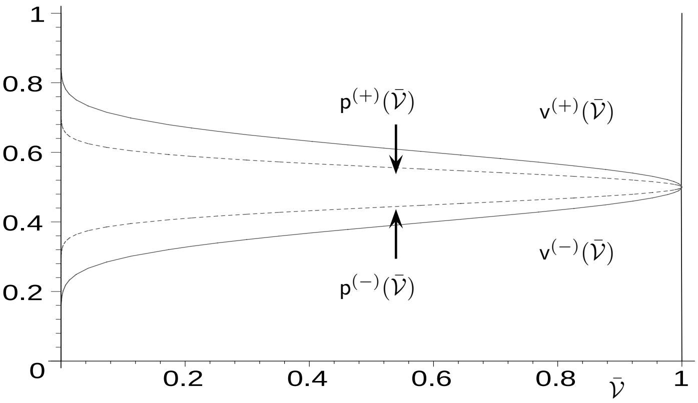


Figure 4: The probabilities $\mathsf{p}^{(\pm)}(\bar{\mathcal{V}})$ and the volume ratios $\mathsf{v}^{(\pm)}(\bar{\mathcal{V}})$, for the dog-flea model, plotted against $\bar{\mathcal{V}}$ for a population of $N = 100$ fleas.

structure imposed by the numbering of the fleas. Each point in $\mathcal{V}(n)$ corresponds to exactly one distribution of the numbered fleas. This restricts the number of points in $\mathcal{V}(n \pm 1)$ which are accessible. It still has the same monotonically decreasing character as that shown in the SRW model and we would expect that, if the dogs begin their day with most fleas on one dog, then as the flea trainer carries out his task, the macrostate volume, and hence the entropy, will increase, rapidly at first, until they are near to their maximum values when they will be subject to small fluctuations. The expected numbers of fleas on Plus and Minus at time $t+1$, given that there are $N^{(\pm)}(t)$ at time t, are

$$\begin{aligned}\langle N^{(\pm)}(t+1)\rangle &= \pi(n, n\pm 1)\left[N^{(\pm)}(t) - \frac{n}{|n|}\right] + \pi(n, n\mp 1)\left[N^{(\pm)}(t) + \frac{n}{|n|}\right] \\ &= N^{(\pm)}(t)\left[1 - \frac{1}{N}\right] + N^{(\mp)}(t)\frac{1}{N}. \qquad (18)\end{aligned}$$

Replacing $N^{(\pm)}(t)$ by their expected numbers gives

$$\langle N^{(\pm)}(t+1)\rangle = \langle N^{(\pm)}(t)\rangle\left[1 - \frac{1}{N}\right] + \langle N^{(\mp)}(t)\rangle\frac{1}{N}. \qquad (19)$$

which have the solution

$$\langle N^{(\pm)}(t)\rangle = \frac{1}{2} + \left(1 - \frac{2}{N}\right)^t \left[N^{(\pm)}(0) - \frac{1}{2}N\right]. \qquad (20)$$

Figs. 5 and 6 show simulations, together with results derived from (20), with 100 fleas beginning with 93 on Minus. In Fig. 5 we see that the numbers of fleas on the two dogs converge towards 50, where they exhibit small fluctuations about the expected values given by (19). Fig. 6 shows the behaviour of the scaled entropy and of its value derived from (19). The probability $\rho(n;t)$ that

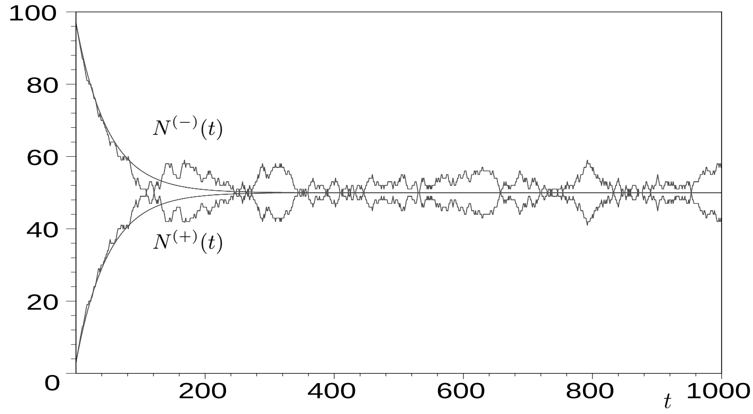


Figure 5: The numbers of fleas $N^{(\pm)}(t)$ on the dogs Plus and Minus plotted against time t for a population of $N = 100$ fleas, beginning with seven fleas on Plus. The smooth curve gives the values of $\langle N^{(\pm)}(t)\rangle$ derived from (20).

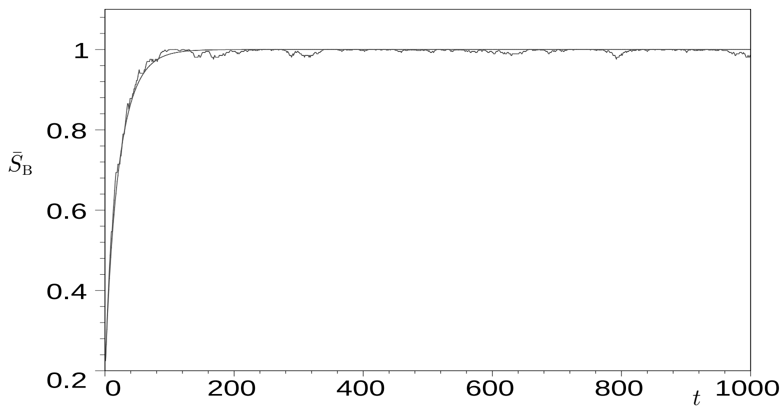


Figure 6: The scaled entropy $\bar{S}_{\rm B}$, for the dog-flea model, plotted against time t for a population of $N = 100$ fleas, beginning with seven fleas on Plus. The smooth curve gives the value derived from (20).

$x(t) = n$ again satisfies (13), except that now the transition probabilities are given by (16). It is not difficult to show that this equation is now satisfied by the equilibrium distribution

$$\rho^\star(n) = \frac{\mathcal{V}(n)}{\displaystyle\sum_{n=-N/2}^{n=N/2} \mathcal{V}(n)} = \frac{1}{2^N}\,\frac{N!}{\left(\frac{1}{2}N+n\right)!\left(\frac{1}{2}N-n\right)!}. \tag{21}$$

For this distribution a number of significant results can be established (see Emch and Liu, 2002, p. 106–112):

(i)

$$\begin{aligned} \mathsf{Prob}[x(t) = n|x(t+1) = k] &= \pi(n,k) \\ &= \mathsf{Prob}[x(t+1) = k|x(t) = n]. \end{aligned} \tag{22}$$

The backward transition probability is equal to the forward transition probability, which is a form of statistical time-reversibility.

(ii) With

$$\Pi(k,n,m) = \mathsf{Prob}[x(t-1) = k \text{ and } x(t+1) = m|x(t) = n], \tag{23}$$

$$\begin{aligned} &\Pi(n-1,n,n-1) : \Pi(n-1,n,n+1) : \Pi(n+1:n:n-1) : \Pi(n+1,n,n+1) \\ &= \frac{\frac{1}{2}N+n}{\frac{1}{2}N-n} : 1 : 1 : \frac{\frac{1}{2}N+n}{\frac{1}{2}N-n}. \end{aligned} \tag{24}$$

It follows that, in this equilibrium distribution, the most probable sequences of three states are reversions towards the long-term mean values $n = 0$, $N^{(+)} = N^{(-)}$ given by (20).

(iii) Let $\nu(n;\tau)$ be the probability that a trajectory starting with $x(0) = n$ will return to state n *for the first time* at time τ. We define, for fixed n the mean and variance of the recurrence time

$$\langle\tau\rangle_n = \sum_{\tau=1}^{\infty} \tau\nu(n;\tau) \quad \text{and} \quad \mathrm{Var}_n(\tau) = \sum_{\tau=1}^{\infty} \left\{\langle\tau\rangle_n - \tau\right\}^2 \nu(n;\tau), \tag{25}$$

It was shown by Kac (1947) that

(a) $$\sum_{\tau=1}^{\infty} \nu(n;\tau) = 1, \tag{26}$$

which means that every state has a probability of one of recurring at some time.

(b) $$\langle\tau\rangle_n = \frac{1}{\rho^\star(n)} = \frac{2^N}{\mathcal{V}(n)}. \tag{27}$$

The recurrence times for $n = 0$ are relatively small but for $n = \frac{1}{2}N$ they are large, even for modest values of N.[18]

(c) For large N and $n \approx N/2$ the variance is of the order of the $\langle\tau\rangle_n$. So the large mean recurrence time has limited significance.

[18]For $N = 10$, $\langle\tau\rangle_0 = 256/63$ and $\langle\tau\rangle_5 = 1024$; for $N = 100$ $\langle\tau\rangle_0 \approx 12.56$ and $\langle\tau\rangle_{50} \approx 12.68 \times 10^{29}$.

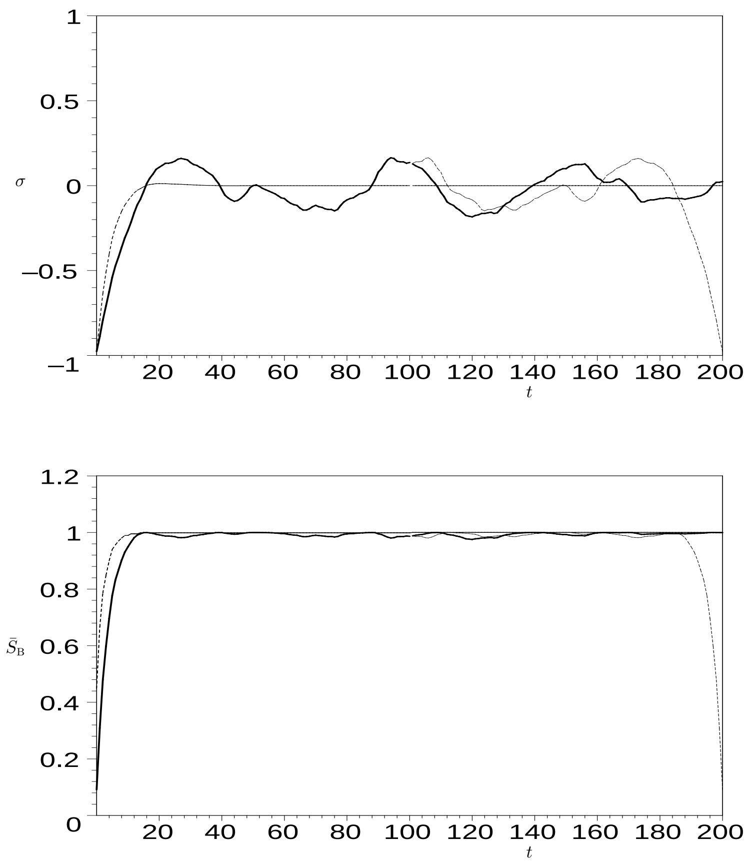


Figure 7: The evolution of the magnetization and the Boltzmann entropy, for the ring model, with $N = 1000$, $m = 56$, $N^{(+)}(0) = 12$. The mean-field value is given by the broken line and, for $t > 100$, the dotted line represents the situation when the rotation is reversed at $t = 100$.

4 The Ring Model

Suppose that we codify the dog-flea model by taking a ring of N sites and place in a random way an up-spin on a site for every flea on Plus and a down-spin on a site for every flea on Minus. The flea jumping is now represented by choosing at random one point equidistant between two sites as a spin-flipper. Then all the spins move one site clockwise, with the one passing through the flipper being flipped. If the location of the spin-flipper is relocated randomly before each rotation the model is exactly equivalent to the dog-flea model. It can be generalized by randomly distributing m spin-flippers; this corresponds to the flea-trainer choosing a team of m fleas with instructions to change dogs.

Now suppose that, having chosen the initial locations of the spin-flippers, we leave them fixed. After the initial distribution of spins and flippers, the model is deterministic. It is reversible simply by rotating the spins in the anticlockwise direction and periodic with period N if m is even and $2N$ if m is odd. This is a version of the Kac ring model.[19] The 'magnetization' at time t is

$$\sigma(t) = \frac{N^{(+)}(t) - N^{(-)}(t)}{N} = \frac{2n(t)}{N} \tag{28}$$

and the Boltzmann entropy is given by the same formula as for the dog-flea model. *Suppose that each spin were equally likely to be flipped.* Then the equations of motion would be

$$\tilde{N}^{(\pm)}(t+1) = \tilde{N}^{(\pm)}(t)\left[1 + \frac{m}{N}\right] + \tilde{N}^{(\mp)}(t)\frac{m}{N} \qquad t = 0, 1, 2, \ldots \tag{29}$$

and it is not difficult to show that

$$\tilde{\sigma}(t) = \sigma(0)\left(1 - \frac{2m}{N}\right)^t. \tag{30}$$

The expectation-value equations (19) for the dog-flea model are just the case $m = 1$ of (29), which would, of course, arise from the dog-flea model if m fleas were instructed to change dogs on each occasion. The (unjustified) equally-likelihood assumption, leads to a 'mean-field theory' where magnetization evolves monotonically from its initial value to zero and the Boltzmann entropy evolves monotonically to its maximum value. As with any mean-field theory fluctuations are smoothed out. Figs. 7 show the evolution of a ring of 1000 spins with 56 spin-flippers including the situation with the rotation reversed at $t = 100$, returning the system to its initial state. (The broken curves are the mirror images of the right-hand half of the curves.) Fig. 8 shows the Boltzmann entropy for the case of 100 spins with 6 spin flippers. These figures clearly show the reversibility and recurrence of this deterministic model. The simplicity of the model allows us to effect reversibility with the consequent fall in entropy in a way which could not be achieved with more 'realistic' systems. The recurrence, for which the time is simply related to the number of spins, again would not occur on realizable time scales for more complicated systems.

5 The Baker's Transformation

This is the transformation, shown in Fig. 9, where a unit square is stretched to twice its width and then cut in half with the right-hand half used to restore the upper half of the unit square. As the mapping ϕ on the cartesian coordinates (x, y) of the unit square it is given by

$$\phi(x,y) = \begin{cases} (2x, \frac{1}{2}y), & \bmod 1, \quad 0 \le x \le \frac{1}{2}, \\ (2x, \frac{1}{2}(y+1)), & \bmod 1, \quad \frac{1}{2} \le x \le 1. \end{cases} \tag{31}$$

A convenient way of writing this transformation is to express x and y as binary strings:

$$\begin{aligned} x &= 0 \cdot x_1 x_2 x_3 \ldots, \\ y &= 0 \cdot y_1 y_2 y_3 \ldots, \end{aligned} \tag{32}$$

[19] See Dresden (1962), Thompson (1972), Bricmont(1995,2001) and Edens (2001) for further discussion of the model and its variants.

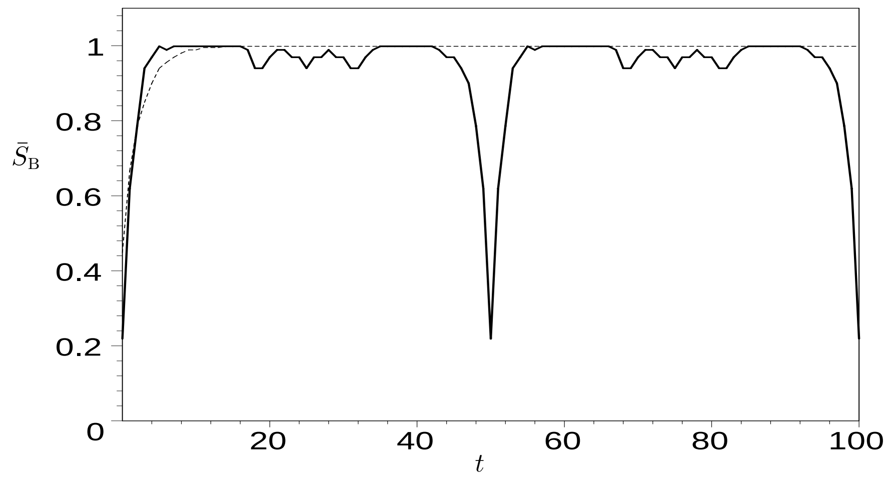


Figure 8: The evolution of the Boltzmann entropy, for the ring model, with $N = 50$, $m = 6$, $N^{(+)}(0) = 2$. The mean-field value is given by the broken line.

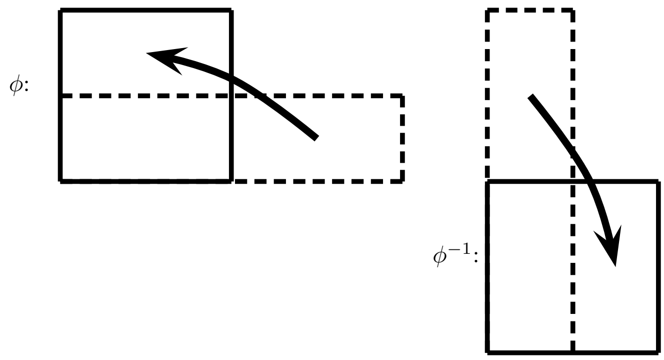


Figure 9: The baker's transformation and its inverse.

where x_j and y_j take the values 0 or 1. Then the baker's transformation takes the form

$$\begin{aligned}&\phi(0\cdot x_1x_2x_3\ldots,0\cdot y_1y_2y_3\ldots)=(0\cdot x_2x_3x_4\ldots,0\cdot x_1y_1y_2\ldots),\\ &\text{with}\\ &\phi^{-1}(0\cdot x_1x_2x_3\ldots,0\cdot y_1y_2y_3\ldots)=(0\cdot y_1x_1x_2\ldots,0\cdot y_2y_3y_4\ldots).\end{aligned} \tag{33}$$

It is clear that the mapping is reversible with $\phi^{-1} = J\phi\, J$ and $J(x, y) = (y, x)$. It can also be shown (Lasota and Mackey, 1994, p. 54–56) to be volume-preserving

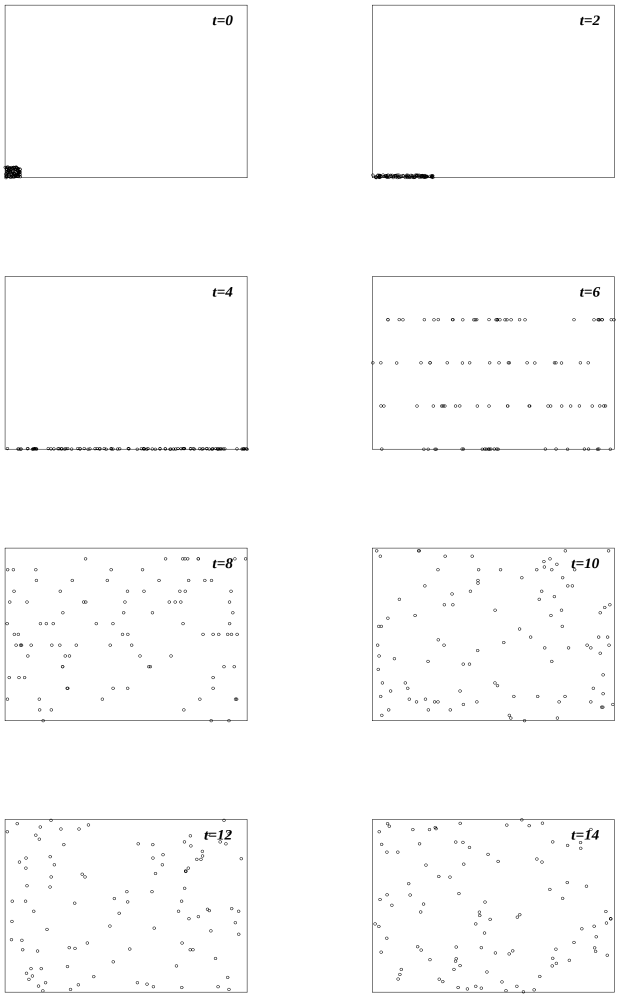


Figure 10: A gas of $N = 100$ particles moving under the baker's transformation.

and thus that the Poincaré (1890) recurrence theorem applies. One way of representing the transformation is to write the bits for the initial point as

$$\ldots y_5 y_4 y_3 y_2 y_1 | x_1 x_2 x_3 x_4 x_5 \ldots . \tag{34}$$

Then ϕ corresponds to moving the vertical bar one step to the right. Now suppose that a trajectory starts at a randomly chosen point in the small square γ given by $0 \leq x < 2^{-m}$, $0 \leq y < 2^{-m}$. This simply means that in (34) there are m entries of zero on each side of the bar. The trajectory will return to γ when, after some translations of the bar, this again happens. We now calculate the mean recurrence time to γ. Let $\mathsf{f}(n)$ be the probability that the point is in γ for the first time (following the starting value) after n steps and let $\mathsf{g}(n)$

be the probability that after n steps the phase point is in γ irrespective of the intermediate values. Then, with $\mathsf{g}(0) = 1$,

$$\mathsf{g}(n) = \sum_{k=1}^{n} \mathsf{f}(k)\mathsf{g}(n-k) \tag{35}$$

Multiplying by z^n and summing over n gives

$$\mathsf{G}(z) = \mathsf{F}(z) + \mathsf{F}(z)\mathsf{G}(z), \tag{36}$$

where

$$\mathsf{F}(z) = \sum_{n=1}^{\infty} \mathsf{f}(n) z^n, \qquad \mathsf{G}(z) = \sum_{n=1}^{\infty} \mathsf{g}(n) z^n. \tag{37}$$

Now

$$\mathsf{g}(n) = \begin{cases} \frac{1}{2^n}, & 1 \le n \le 2m, \\ \frac{1}{2^{2m}}, & 2m \le n. \end{cases} \tag{38}$$

So

$$\mathsf{G}(z) = \frac{2^{2m}(1-z) + z^{2m+1}}{2^{2m}(2-z)(1-z)} \tag{39}$$

and, from (36),

$$\mathsf{F}(z) = \frac{2^{2m}(1-z) + z^{2m+1}}{2^{2m}(3-z)(1-z) + z^{2m+1}}. \tag{40}$$

Of course, $\mathsf{F}(1) = 1$ and $\mathsf{F}'(1) = 2^{2m}$ is the mean recurrence time to γ. In the hierarchy of special dynamic properties: ergodic, mixing, Kolmogorov, Bernoulli, each implies the one preceding it. The baker's transformation is Bernoulli.[20] To show this divide the unit square into the partition $B_2 = \{b_0, b_1\}$, where b_0 is the set where $0 \le x < \frac{1}{2}$ and b_1 the set where $\frac{1}{2} \le x < 1$. Then (x, y) will be in b_{x_1}. The complete record of the labels of a trajectory is the string (34) (without the bar). This is distinct for each trajectory of points and the members are uncorrelated so the partition is Bernoulli.

A rather more interesting application of the baker's transformation is to consider a 'gas' of N points. Then we can start all the points in some small subset of the unit square and watch them evolve. Suppose the square is divided into 2^{2m} equal square cells and all the particles begin in the bottom left-hand cell. Fig. 10 shows the evolution for a gas of $N = 100$ particles with $m = 4$ (256 cells). Now we suppose that the macrostates correspond to identifying the number of particles N_{ij} in each of the cells $i, j = 1, 2, \ldots, 2^m$. Then

$$\mathcal{V}(\{N_{ij}\}) = \frac{N!}{\prod_{i,j} N_{ij}!} \frac{1}{2^{2m}}, \qquad \mathcal{V}_{\min} = \frac{1}{2^{2m}}. \tag{41}$$

Fig. 11 shows the evolution of the Boltzmann entropy (scales as usual by it

[20] A system is Bernoulli if there is a Bernoulli-partition B_n of the invariant space Σ of the transformation. of Σ. This is defined in the following way. Let B_n be a partition of the Σ into n subsets. We label the members of B_n with the integers $[0, 1, \ldots, n-1]$ and, for any trajectory and some $\triangle t$, record the infinite sequence of numbers $\mathcal{S} = \{\ldots, s_{-2}, s_{-1}, s_0, s_1, s_2, \ldots\}$, where the trajectory is in the set b_{s_k} labelled s_k at time $k\triangle t$. The partition is Bernoulli if all sequences $\mathcal{S}$ are uncorrelated and no two distinct trajectories have the same sequence.

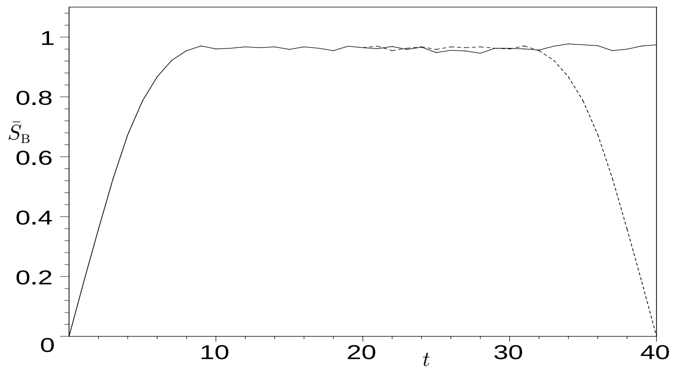


Figure 11: The evolution of the Boltzmann entropy of $N = 100$ particles moving under the baker's transformation. The broken line shows the return to the initial state when the transformation $J(x, y) = (y, x)$ is applied at $t = 20$.

maximum value). The entropy will return to its initial value if all the particles arrive in the same cell. The mean time for this to occur can be calculated by the same procedure given above but with (38) replaced by

$$\mathsf{g}(n) = \begin{cases} \frac{1}{2^{n(N-1)}}, & 1 \le n \le 2m, \\ \frac{1}{2^{2m(N-1)}}, & 2m \le n. \end{cases} \tag{42}$$

This then yields the result that the mean time for the entropy to return to its initial value is $2^{2m(N-1)}$ steps. For $m = 4$, $N = 100$ this approximates to 10^{238} steps.

6 The Spin–Echo System

Consider the simple model in which a magnetic dipole of moment $\boldsymbol{\mu}$ is fixed at its centre but is free to rotate in the presence of a constant magnetic field $\boldsymbol{B}$. The equation of motion of the dipole will be

$$\dot{\boldsymbol{\mu}}(t) = g\,\boldsymbol{\mu}(t) \wedge \boldsymbol{B}, \tag{43}$$

where g is the gyromagnetic ratio. It is not difficult to show that motion of the dipole is a precession at a constant angle to $\boldsymbol{B}$. In particular, if $\boldsymbol{\mu}$ is located at the origin of a cartesian coordinate system with $\boldsymbol{B}$ in the direction of the negative z-axis and if initially $\boldsymbol{\mu}$ lies in the $x - y$ plane, its subsequent motion remains in the $x - y$ plane and is given by

$$\boldsymbol{\mu}(t) = (\mu\cos(\theta(t)), \mu\sin(\theta(t))), \tag{44}$$

where

$$\theta(t) = \theta(0) + \omega t, \qquad \omega = B\,g. \tag{45}$$

Suppose that at some time $t = \tau$ the magnetic field $\boldsymbol{B}$ is turned off and a field $\boldsymbol{B}'$, in the direction of the x–axis is turned on for a time $t' = \pi/B'g$. The effect of this will be to rotate the dipole through an angle π about the x-axis, translating it position from $\theta(\tau) = \theta(0) + \omega\tau$ to $\theta'(\tau) = 2\pi - \theta(0) - \omega\tau$, a reflection in the x-axis. Then after a further time[21] τ

$$\theta(2\tau) = \theta'(\tau) + \omega\tau = 2\pi - \theta(0). \tag{46}$$

If the dipole begins pointing along the $x-$ axis then it returns to its initial position (the *echo effect*). There is nothing remarkable about this result. A similar effect could have been produced by reversing the direction of rotation. However, we are interested, as in our previous examples, in an assembly of microsystems. Consider the collection $\boldsymbol{\mu}_i$, $i = 1, 2, \ldots, N$ of such dipoles with angular velocities ω_i in the range $[\omega_{\min}, \omega_{\max}]$ and plot their evolutions in the $\theta - \omega$ plane. Suppose that, $N = 500$, $\tau = 100$, $\omega_{\min} = 0.75$ and $\omega_{\max} = 1.25$ and that the ω_i are chosen randomly in $[\omega_{\min}, \omega_{\max}]$ with $\theta_0 = 0$ for all the dipoles. Then we have the situation shown in Fig. 12. The 'gas' of dipole phase points spreads into the single–dipole phase space Γ_1. Since the angular velocities have been chosen randomly the assembly is quasi-periodic. The system is Hamiltonian and will satisfy the Poincaré recurrence theorem. For 'most' initial points, if the position is not reflected at some time τ, the phase point $(\boldsymbol{\theta}, \boldsymbol{\omega}) = (\theta_1, \ldots, \theta_N, \omega_1, \ldots, \omega_N)$ in the $2N$-dimensional phase space Γ_N nevertheless returns to within a neighbourhood of its initial value.[22] Of course, if the initial angular velocities are chosen to be commensurate, the system will be periodic and will return exactly to its initial point.

There would be nothing particularly special about this model, if it were not for the fact that it has been realized experimentally. Although the system is not cooperative, the realization of a controlled return to its initial state for a many-component system is a considerable achievement. Hahn (1950) used a sample of glycerin in a magnetic field. By manipulating the components of the magnetic field he was able to start with the dipole moments of the protons in the x–direction, make them precess around the z–axis and then reflect the directions of the dipoles in the x–axis to achieve the echo effect with the dipoles returning to their initial alignment.[23] This system has aroused some interest in relation to questions of reversibility in statistical mechanics (Ridderbos and Redhead, 1998; Ridderbos, 2002). Here our purpose is not so much to enter these disputes as to provide simulations which could lead to clarification of the argument. We shall calculate not only the Boltzmann entropy but the fine-grained and course-grained versions of the Gibbs entropy.

As in the case of the baker's transformation, described in Sec. 5, the Boltzmann entropy is obtained by defining macrostates by course-graining. We divide the single–dipole phase space Γ_1 into $n_\theta \times n_\omega$ rectangular cells with edges parallel to the θ and ω axes and of lengths $\triangle\theta = 2\pi/n_\theta$ and $\triangle\omega = (\omega_{\max} - \omega_{\min})/n_\omega$ respectively. We label the cells by the indices $\eta_\theta = 1, 2, \ldots, n_\theta$, $\eta_\omega = 1, 2, \ldots, n_\omega$. For a particular phase point $(\boldsymbol{\theta}, \boldsymbol{\omega})$, in Γ_N, let $N(\eta_\theta, \eta_\omega; \boldsymbol{\theta}, \boldsymbol{\omega})$ be the number of dipoles with phase points in Γ_1 in the cell $(\eta_\theta, \eta_\omega)$. We define the macrostate $\mathcal{S}(\boldsymbol{\theta}, \boldsymbol{\omega})$ as the set of points $(\boldsymbol{\theta}', \boldsymbol{\omega}')$ such that

$$N(\eta_\theta, \eta_\omega; \boldsymbol{\theta}', \boldsymbol{\omega}') = N(\eta_\theta, \eta_\omega; \boldsymbol{\theta}, \boldsymbol{\omega}), \qquad \eta_\theta = 1, 2, \ldots, n_\theta, \;\; \eta_\omega = 1, 2, \ldots, n_\omega$$

[21] We neglect the time t' needed to produce the reflection.

[22] The recurrence time will, of course, be dependent on the size of the neighbourhood.

[23] The variations in the angular velocities were achieved from small variations in the strength of the magnetic field throughout the sample.

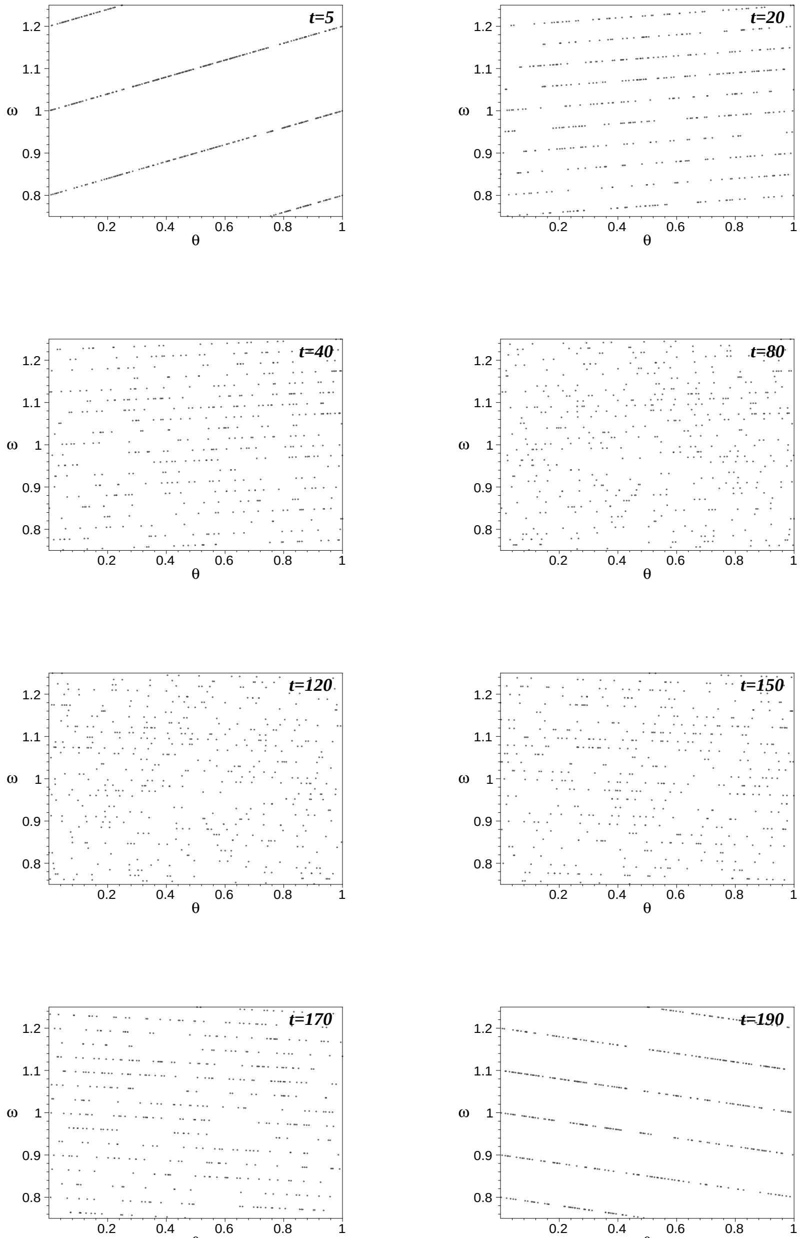


Figure 12: An assembly of $N = 500$ rotating dipoles.

$$ \tag{47} $$

and thus

$$\mathcal{V}(\boldsymbol{\theta}, \boldsymbol{\omega}) = \Omega(\boldsymbol{\theta}, \boldsymbol{\omega})[\triangle\theta \, \triangle\omega]^N, \qquad \mathcal{V}_{\min} = [\triangle\theta \, \triangle\omega]^N, \tag{48}$$

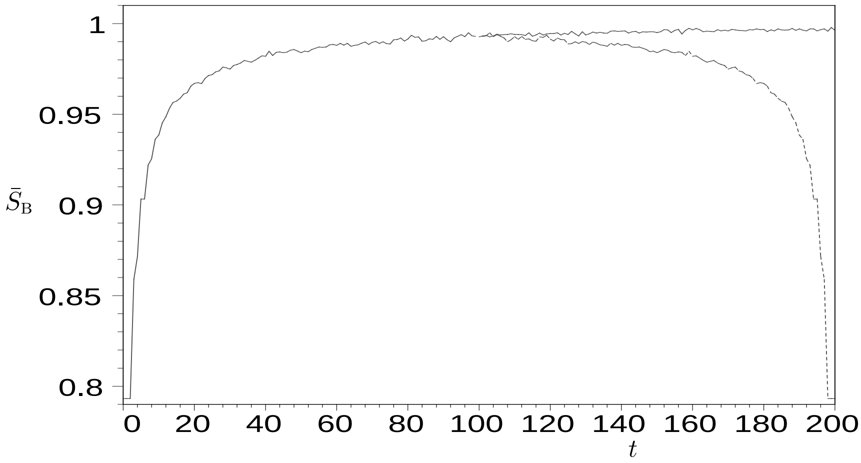


Figure 13: The evolution of the Boltzmann entropy of the dipole assembly. After $t = \tau = 100$ the broken line gives the echo.

where

$$\Omega(\boldsymbol{\theta},\boldsymbol{\omega}) = \frac{N!}{\prod_{\eta_\theta=1}^{n_\theta}\prod_{\eta_\omega=1}^{n_\omega}[N(\eta_\theta,\eta_\omega;\boldsymbol{\theta},\boldsymbol{\omega})]!} \tag{49}$$

From (1),

$$S_{\mathrm{B}}(\boldsymbol{\theta},\boldsymbol{\omega}) = k_{\mathrm{B}}\ln[\Omega(\boldsymbol{\theta},\boldsymbol{\omega})], \tag{50}$$

The maximum value of S_{B} is given when each cell is occupied by an equal number of dipoles and in Fig. 13 we show the Boltzmann entropy, with $n_\theta = n_\omega = 100$, scaled by its maximum value, for the same evolution that is shown in Fig. 12. The continuous and broken lines for $t > 100$ correspond respectively to the evolutions without and with the echo-effect.

We now calculate the fine-grained Gibbs entropy. Suppose that the initial probability density function is concentrated and uniform over the interval $\theta \in [0,\theta_0]$ ($\theta_0 < 2\pi$). Then

$$\rho_1(\theta,\omega;0) = \frac{\mathcal{H}(\theta) - \mathcal{H}(\theta-\theta_0)}{\theta_0(\omega_{\mathrm{max}} - \omega_{\mathrm{min}})}, \tag{51}$$

where $\mathcal{H}(\theta)$ is the *Heaviside unit function.* Now the periodicity of (45) can be manifested explicitly in the form

$$\theta(t) = \mathcal{F}[\theta(0) + \omega t], \tag{52}$$

where

$$\mathcal{F}[x] = 2\pi \times \textsf{Non-Integer Part}\left(\frac{x}{2\pi}\right). \tag{53}$$

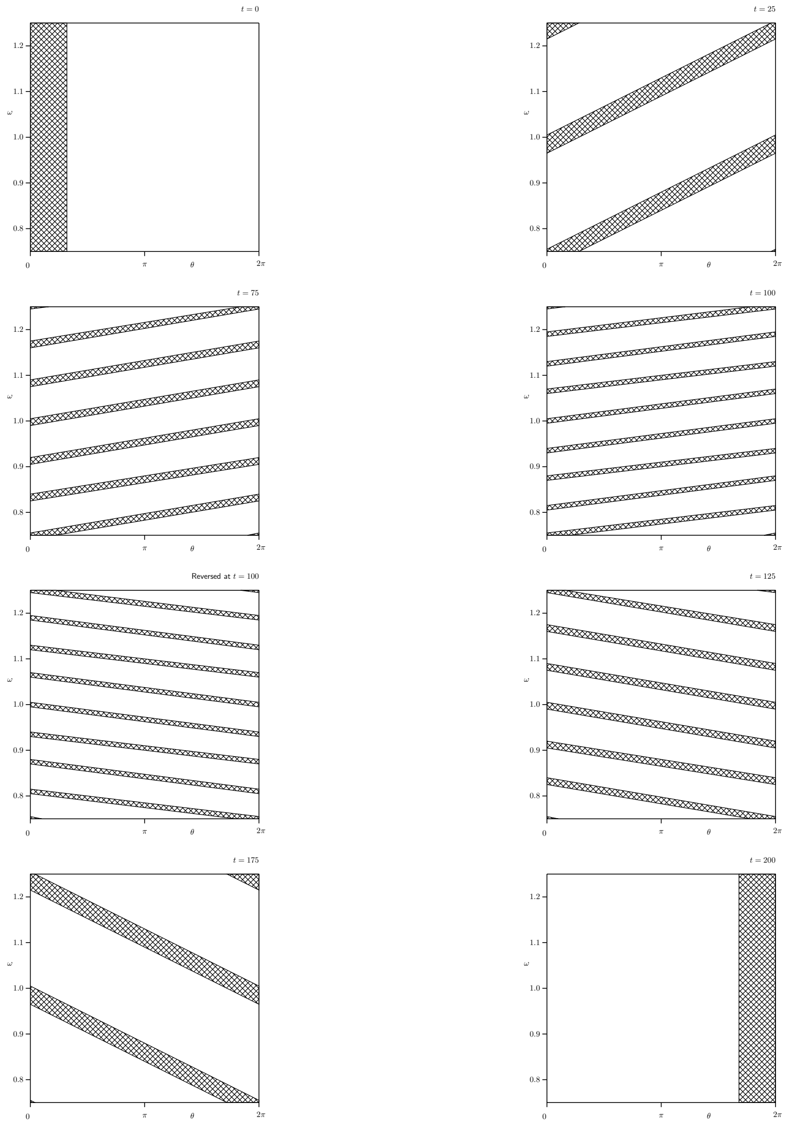


Figure 14: The evolution of the fine-grained probability density function $\rho_1(\theta, \omega; t)$ with the echo occurring at $t = \tau = 100$.

Then

$$\rho_1(\theta, \omega; t) = \begin{cases} \dfrac{\mathcal{H}(\theta - \mathcal{F}[\omega t]) - \mathcal{H}(\theta - \mathcal{F}[\theta_0 + \omega t])}{\theta_0(\omega_{\max} - \omega_{\min})}, & \\ & \mathcal{F}[\omega t] < \mathcal{F}[\theta_0 + \omega t], \\ \dfrac{\mathcal{H}(\theta - \mathcal{F}[\omega t]) - \mathcal{H}(\theta - \mathcal{F}[\theta_0 + \omega t]) + \mathcal{H}(\theta) - \mathcal{H}(\theta - 2\pi)}{\theta_0(\omega_{\max} - \omega_{\min})}, & \\ & \mathcal{F}[\theta_0 + \omega t] < \mathcal{F}[\omega t]. \end{cases} \tag{54}$$

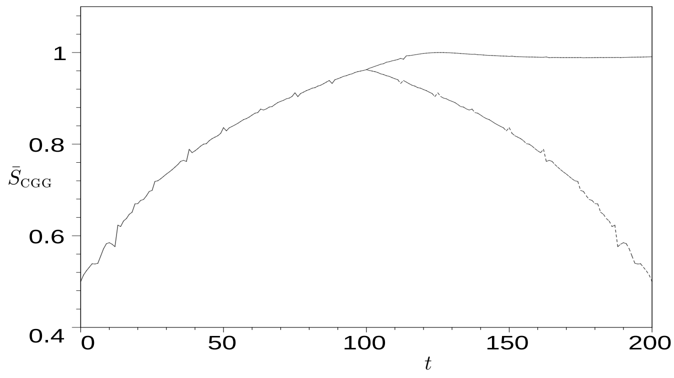


Figure 15: The evolution of the course-grained Gibbs entropy of the dipole assembly. After $t = \tau = 100$ the broken line corresponds to the echo.

If the echo transformation $\theta \to 2\pi - \theta$ is applied at the time τ the one-spin probability density function for $t > \tau$ is given, in terms of (54) by $\rho_1(2\pi - \theta, \omega, 2\tau - t)$. The evolution of this fine-grained probability density function with $\tau = 100$ is shown in Fig. 14.

Substituting from (54) into (5) gives

$$S_{\rm FGG}[\rho_N(t)] = k_{\rm B} N \ln\{\theta_0(\omega_{\rm max} - \omega_{\rm min})\}, \tag{55}$$

This result is simply an expression of the well-known result that the fine-grained Gibbs entropy is invariant with respect to time. The course-grained Gibbs entropy is now calculated using the same course-graining as was used to obtain the macrostates for the Boltzmann entropy. $S_{\rm CGG}[\rho_N(t)]$ will have a maximum value when the hatched area in Fig. 12 is spread evenly over the cells. Then $\bar{\rho}_1(k;t) = (\triangle\theta\triangle\omega)/\{2\pi(\omega_{\rm max} - \omega_{\rm min})\}$. Substituting into (7) (with $\nu = \triangle\theta\triangle\omega$) gives

$$(S_{\rm CGG})_{\rm max} = k_{\rm B} N \ln\{2\pi(\omega_{\rm max} - \omega_{\rm min})\}. \tag{56}$$

We adopt the strategy (i) of Sec. 1.2 and course-grain the fine-grained density as time evolves (rather than performing successive re-course-grainings). The results for $n_\theta = n_\omega = 100$ are shown in Fig. 15. Ridderbos and Redhead (1998) have shown[24] that the course-grained entropy tends to its maximum value (56) as $t \to \infty$ and our simulations in Fig. 15 support this result.

7 Some Comments

The purpose of this work has been to provide results for simulations of a number of simple models which have relevance to statistical mechanics. In Sec. 1 we commented that approximations used to solve statistical mechanical models,

[24]Thus proving the system is mixing.

while often practically very useful, tend to vitiate inferences of a more fundamental kind. In general, of course, computational work suffers from the same defects and in the light of this we should consider the weight which can be given to our results. The most serious problem would be if temporal iteration led to an accumulation of numerical errors. This will not be the case for the models of Secs. 2–4, where the phase space is discrete and the volumes of the macrostates are integer. The only approximation here is in the calculation of the entropy for the purpose of graphical representation. For both the spin–echo system and baker's transformation the phase space is a subset of $\mathbb{R}^2$. The iteration of the spin–echo system is given by (45) and since all the variables are real there will be some numerical rounding which will also affect the calculation of the number of phase-points in the cells for the Boltzmann entropy and the course-grained densities for the course-grained Gibbs entropy. However, our results would indicate that using MAPLE with hardware accuracy and with the level of course-graining used the results are not significantly compromised.[25]

The most interesting case is the baker's transformation. The initial state of a microsystem is given by a binary string like (34). Each iteration is equivalent to moving the bar one step to the left and the cell in which the microsystem is situated is given by the m bits each side of the bar. For enough information to be contained in the initial string so that the cell for the the j–th iteration is predictable the binary strings for the cartesian coordinates must be accurate to $j + m$ places. This level of accuracy will also give the maximum number of iterations which can be performed with the possibility of reversing the direction and returning with certainty to the initial cell. This latter observation provides a good computational test. We found for our system that with $m = 4$ we could perform around twenty iterations and return the system along the same set of cells. This is shown if Fig. 11. Of course, the unreversed part of the curve for $t > 20$ will rely on bits which were not contained in the initial point.[26]

There is much discussion in the literature of the models treated in this work; in particular in relation to rival viewpoints about the appropriate form for non-equilibrium entropy. However, because of the analytic difficulty, there is a lack of hard information about the way these behave over time. Of course, our results do not *prove* anything about that. But, we suggest that they provide evidence that *typically* the Boltzmann entropy behaves in the way suggested in Sec. 1 and that the behaviour of the course-grained Gibbs entropy can be described in a similar way.

References

Bogoliubov, N. N. (1962). Problems of a dynamical theory in statistical physics, *in* J. de Boer and G. E. Uhlenbeck (eds), *Studies in Statistical Mechanics, Vol. 1*, North-Holland, pp. 1–118.

Bricmont, J. (1995). Science of chaos or chaos in science?, *Physicalia* **17**: 159–208.

[25]We are able to produce an echo with entropy identical to the initial value to within visual accuracy.

[26]But they are qualitatively indistinguishable from those which, in principle could have been, if the accuracy had been greater.

Bricmont, J. (2001). Bayes, Boltzmann and Bohm: probabilities in physics., *in* J. Bricmont, D. Dürr, M. C. Galvotti, G. Ghirardi, F. Petruccione and N. Zanghi (eds), *Chance in Physics: Foundations and Perspectives*, Springer, pp. 3–21.

Drazin, P. G. (1992). *Nonlinear Systems*, Cambridge U. P.

Dresden, M. (1962). A study of models in non-equilibrium statistical mechanics, *in* J. de Boer and G. E. Uhlenbeck (eds), *Studies in Statistical Mechanics, Vol. 1*, North-Holland, pp. 303–343.

Edens, B. (2001). *Semigroups and symmetry: an investigation of Prigogine's theories*, PhD thesis, Institute for History and Foundations of Science, Utrecht University. Available at:
`http://philsci-archive.pitt.edu/documents/disk0/00/00/04/36/`.

Ehrenfest, P. and Ehrenfest-Afanassjewa, T. (1907). Ueber zwei bekannte Einwände gegen das Boltzmannshe H-Theorem, *Phys. Zeitschrift* **8**: 311–314.

Ehrenfest, P. and Ehrenfest-Afanassjewa, T. (1912). *The Conceptual Foundations of the Statistical Approach in Mechanics*, English translation, Cornell U. P.

Emch, G. and Liu, C. (2002). *The Logic of Thermostatistical Physics*, Springer.

Gibbs, J. W. (1902). *Elementary Principles in Statistical Mechanics*, Yale U. P. Reprinted by Dover Publications, 1960.

Hahn, E. L. (1950). Spin echoes, *Phys. Rev.* **80**: 580–594.

Kac, M. (1947). On the notion of recurrence in discrete stochastic processes., *Bull. Amer. Math. Soc.* **53**: 1002–1010.

Kac, M. (1959). *Probability and Related Topics in the Physical Sciences*, Interscience.

Lasota, A. L. and Mackey, M. C. (1994). *Chaos, Fractals and Noise*, Springer.

Lavis, D. A. and Bell, G. M. (1999). *Statistical Mechanics of Lattice Systems, Vols. 1 and 2*, Springer.

Lebowitz, J. L. (1993). Boltzmann's entropy and time's arrow, *Physics Today* **46**: 32–38.

Parzen, E. (1962). *Stochastic Processes*, Holden-Day.

Poincaré, H. (1890). On the three-body problem and the equations of dynamics, *in* S. G. Brush (ed.), *Kinetic Theory Vol. 2 (1966)*, Pergamon, pp. 194–202. English Reprint.

Ridderbos, K. (2002). The course-graining approach to statistical mechanics: how blissful is our ignorance?, *Stud. Hist. Phil. Mod. Phys.* **33**: 65–77.

Ridderbos, T. M. and Redhead, M. L. G. (1998). The spin-echo experiments and the second law of thermodynamics, *Foundations of Physics* **28**: 1237–1270.

Thompson, C. J. (1972). *Mathematical Statistical Mechanics*, Princeton U. P.