MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A.I. Memo No. 1346 September 1992

Time-Reversible Maxwell's Demon

P. A. Skordos
Center for Nonlinear Studies
Los Alamos National Lab, B258, Los Alamos, NM 87545
and
Massachusetts Institute of Technology
545 Technology Square, NE43-432, Cambridge, MA 02139

Abstract

A time-reversible Maxwell's demon is demonstrated which creates a density difference between two chambers initialized to have equal density. The density difference is estimated theoretically and confirmed by computer simulations. It is found that the reversible Maxwell's demon compresses phase space volume even though its dynamics are time reversible. The significance of phase space volume compression in operating a microscopic heat engine is also discussed.

This report describes research done at Los Alamos National Laboratory, Los Alamos NM and at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the MIT laboratory's research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-89-J-3202 and by the National Science Foundation under grant number MIP-9001651.

1 Introduction

The second law of thermodynamics is usually attributed to the fact that states of maximum disorder in a statistical system have the largest probability of occurrence among all possible states. The exact form of the microscopic dynamics is usually assumed to play a secondary role. Yet the microscopic dynamics can not be arbitrary. There are certain conditions that the microscopic dynamics of a system must satisfy for the second law to hold. This paper examines a system of hard core disks whose microscopic dynamics do not satisfy the second law and can be viewed as a time-reversible Maxwell's demon.

The hard core disks move and collide with each other elastically, and are kept inside a container that is separated into two chambers by a membrane. The membrane interacts with the disks by transforming an incident disk's velocity according to a time-reversible and energy conserving rule to be described below. The interaction between the disks and the membrane makes the system of disks non-ergodic even though the dynamics are chaotic, energy conserving, and time reversible.

The next section describes in detail the membrane and disks system and explains how the membrane interaction leads to a density difference between the two chambers. Section 3 estimates the density difference theoretically,

and section 4 presents the results of computer simulations confirming the theoretical estimate. Section 5 discusses the dynamics of the membrane and disks system in a multidimensional phase space, and reviews a few other examples of dynamical systems that exhibit non-ergodic behavior.

Section 6 shows that the membrane interaction compresses phase space volume when a disk penetrates the membrane. It also demonstrates that if the membrane interaction obeyed incompressible dynamics, then it could not create a density difference (for piecewise differentiable maps). Thus the loss of ergodicity in the membrane and disks system must be attributed to compression of phase space volume. Section 7 shows that a heat engine that consists of a microscopic rectifier and a cooling mechanism also compresses phase space volume, but in this case the compression of phase space volume is caused by external forces.

Section 8 looks at the membrane system as a Maxwell's demon whose dynamics are time-reversible, and compares it to the traditional Maxwell's demon whose dynamics are irreversible. The comparison suggests that compression of phase space volume is more general than irreversibility for reducing the entropy of an isolated system of disks. An information theoretic model of the new demon is also discussed.

2 The System

The system of membrane and disks (shown in figure 1) consists of a box containing hard core disks, and a membrane that separates the box into two chambers of equal volume. The hard core disks move and collide with each other elastically. The membrane interacts with every incident disk according to a rule that either reflects the disk back, or allows the disk to penetrate the membrane while permuting and negating the disk's velocity components according to the following equation.

$$\begin{aligned}
+45: &\left\{\begin{array}{rcl} V_1' &=& V_2 \\ V_2' &=& V_1 \end{array}\right\} \quad \text{if} \quad \left\{\begin{array}{l} (V_1<0 \quad V_2<0 \quad |V_1|>|V_2|) \text{ or} \\ (V_1>0 \quad V_2>0 \quad |V_1|<|V_2|) \end{array}\right\} \\
-45: &\left\{\begin{array}{rcl} V_1' &=& -V_2 \\ V_2' &=& -V_1 \end{array}\right\} \quad \text{if} \quad \left\{\begin{array}{l} (V_1<0 \quad V_2>0 \quad |V_1|>|V_2|) \text{ or} \\ (V_1>0 \quad V_2<0 \quad |V_1|<|V_2|) \end{array}\right\} \\
+90: &\left\{\begin{array}{rcl} V_1' &=& -V_1 \\ V_2' &=& V_2 \end{array}\right\} \quad \text{otherwise}
\end{aligned} \tag{1}$$

The labels +45, −45, and +90 are motivated by the discussion of section 8 that views the membrane as a time-reversible Maxwell's demon. The above equation says that an incident disk in octants 2,4,5, and 7 (refer to figure 2) reverses the x-component of its velocity when it hits the membrane and is not allowed to penetrate. It also says that an incident disk in the remaining octants penetrates the membrane by being deflected toward the x-axis when

coming from the left, and away from the x-axis when coming from the right.

Figure 2 illustrates graphically equation 1. The vertical solid line in figure 2 denotes the membrane, and the other solid lines denote a division of the plane into octants. The two dashed lines inside the first and sixth octants (counting counterclockwise) denote trajectories that are deflected and penetrate the membrane according to the case +45 of equation 1. The transformation of velocities takes place instantaneously when the center of an incident disk reaches the membrane.

Another way of examining equation 1 is to rewrite the equation as a map of the velocity angle (impact angle). This is shown in figure 3 where Θ is the impact velocity angle and Θ' is the transformed velocity angle after the membrane interaction has occurred. Both Θ and Θ' range from $-\pi$ to π. Note that a completely transparent membrane would be a straight line of slope "1" passing through the origin, and a completely impenetrable membrane (mirror-reflecting) would be a line of slope "-1" shifted π radians from the origin and wrapped around periodically. Figure 3 shows that the map of equation 1 consists of eight line segments. The eight line segments can be viewed as the result of breaking a mirror-reflecting line of slope "-1" and shifting the broken line segments that are shown dashed in figure 3 up and down. The dashed line segments of figure 3 correspond to impact angles

that penetrate the membrane, while the solid line segments correspond to impact angles that are mirror-reflected back.

The membrane interaction described by equation 1 leads to a large density difference between the two chambers. This is accomplished by exploiting the thermal motion of the disks and the impact rate of disks hitting the membrane. Assuming that the velocities of the disks are distributed isotropically inside the container, it follows from geometrical considerations that the impact rate of disks hitting the membrane must be a cosine of the impact angle in absolute value,

$$\text{impact-rate} \quad \propto \quad |\cos\Theta| \ . \tag{2}$$

The membrane exploits this cosine distribution of impact angles by allowing disks with "high rate" impact angle to penetrate from the right, and allowing disks with "low rate" impact angle to penetrate from the left. To achieve reversibility, the remaining impact angles ("low rate" from the right and "high rate" from the left) are blocked and do not penetrate the membrane. They are simply reflected back. Furthermore, each "high rate" angle from the right is rotated onto a "low rate" angle when the disk penetrates the membrane, and vice verca.

The membrane interaction of equation 1 makes the membrane more permeable from the right side than from the left. This leads to an excess flux of

disks from the right side of the membrane, and creates a density difference between the two chambers. When the density difference reaches an equilibrium value (which is calculated in the next section), the fluxes of disks between the two sides of the membrane become equal.

3 Estimate of the Density Difference

At equilibrium the fluxes of disks from the left and from the right side of the membrane are equal to each other. In other words if N_L is the normalized density in the left chamber, and $P_{(L\to R)}$ is the probability of an *individual disk* to penetrate the membrane coming from the left, we require that

$$N_L \; P_{(L\to R)} \;=\; N_R \; P_{(R\to L)} \quad . \tag{3}$$

We can estimate $P_{(L\to R)}$ using the fact that among all the disks that strike the membrane from the left only those with trajectories in the third and sixth octants of figure 2 are allowed to penetrate. In particular, these disks have impact angles in the intervals $(\pi/4, \pi/2)$ and $(4\pi/4, 5\pi/4)$. In addition, we assume that the probability of an individual disk to strike the membrane varies as the cosine of the impact angle and is independent of the density in each chamber. Although the total impact rate depends linearly on the density of the disks, the probability of an individual disk to reach the membrane is

independent of the density to a first approximation. Thus we get,

$$P_{(L\to R)} \simeq 2\ C \int_{\pi/4}^{\pi/2} \cos\theta\, d\theta \simeq 2\ C\ 0.3\ , \tag{4}$$

for some normalization constant C; and similarly,

$$P_{(R\to L)} \simeq 2\ C \int_{4\pi/4}^{5\pi/4} \cos\theta\, d\theta \simeq 2\ C\ 0.7\ , \tag{5}$$

which gives

$$N_L \simeq 0.7 \quad \text{and} \quad N_R \simeq 0.3\ . \tag{6}$$

In other words the system of membrane and disks reaches equilibrium when the fluxes of disks from the left and right side of the membrane are balanced, and this happens when the normalized density is approximately 0.7 in the left chamber and approximately 0.3 in the right chamber.

4 Simulation Results

To check the theoretical results of the previous section, a two-dimensional system of hard core disks with a membrane has been simulated. The computer program used in these simulations is the same program as the one described in detail in reference [1] with a few modifications to simulate the membrane. The following membrane rule has been added to the program: Whenever the center of an incident disk reaches the membrane, the disk's velocity is transformed according to equation 1.

In the experiments reported below forty disks are used. The size of each chamber is $24.3 \times 10^{-13} cm^2$ (equal size chambers), and the disk radius is $3 \times 10^{-8} cm$. These numbers give a mean free path of the order of $10^{-6} cm$ which is approximately the length of each chamber. The average speed of each disk is $3.56 \times 10^4 cm/sec$.

Figure 4 shows the time average of the number of disks in the left chamber, building up from an initial value of 0.5 (normalized) to a value of 0.7 (approximately) as a result of the membrane interaction. The smooth curve displays the cumulative time average of the number of disks in the left chamber, which approaches a steady value as time increases. The noisy curve displays a sequence of short time averages of the number of disks in the left chamber (averaged over $1.25 \times 10^{-10} sec$ intervals) and provides an indication of the density fluctuations for the chosen system parameters.

Figure 5 shows the same quantities as figure 4, but plots them on a much longer time scale. In addition, the running averages are based on longer time intervals ($25 \times 10^{-10} sec$), so the size of fluctuations is accordingly reduced. The cumulative time average of the number of disks in the left chamber is a straight line that intersects the y-axis at the value of 0.68 (normalized). This corresponds to 0.69 density (normalized) if we take into account that the unequal numbers of disks in the two chambers changes the available area

in each chamber. The simulation results are in good agreement with the theoretical estimate of 0.7 normalized density difference in section 3.

5 Ergodicity

The membrane and disks system is now discussed from the perspective of multidimensional phase space. A system of N disks can be represented in R^{4N} phase space as a $4N$ vector $(\ldots, X_i, Y_i, U_i, V_i, \ldots)$ of real numbers. The $4N$ vector is called the representative point of the system, and it specifies exactly the positions and velocities of all particles in the system at any given time (all particles have equal mass). As the system evolves in time, the representative point moves inside a constant energy subsurface of the R^{4N} phase space because the total energy is conserved. The total linear momentum is not conserved because wall collisions reflect a disk's momentum $U_i' = -U_i \quad \text{or} \quad V_i' = -V_i$ and because membrane interactions permute and/or reflect a disk's momentum according to equation 1. The total linear momentum is only conserved in a time average sense, and this has been checked by computer simulations. The energy subsurface that is accessible to the representative point will be denoted by Ω.

The membrane and disks system is not ergodic because it does not spend equal times in equal regions of Ω [2, p.68]. If the representative point of the

system visited regions of Ω that have more disks in the left chamber as often as regions of Ω that have more disks in the right chamber, then the time average density would be equal in the two chambers. Instead, the membrane and disks system approaches irreversibly a state of 0.7/0.3 density difference independent of initial conditions.

In general there are many ways in which a system can fail to be ergodic. A trivial way to lose ergodicity in the context of billiard balls is to remove all interactions between the disks. Then the disks can not see each other and bounce between the walls of the container undisturbed. This system does not attain a Maxwellian velocity distribution. It is a trivial example that shows that collisions between disks are necessary for modeling ideal gas. Binary collisions give rise to chaotic dynamics, and allow a system of hard disks to explore fully its phase space.

On the other hand, binary collisions are not enough by themselves to guarantee ergodicity. A system based on binary collisions and some other dynamics can fail to be ergodic if the additional dynamics introduce an attractor in phase space, which may occur when the evolution map $M : \Omega \rightarrow \Omega$ does not conserve measure in phase space. To conserve measure, the Jacobian determinant of the evolution map M must be unity in absolute value. The simplest example of an evolution map that does not conserve measure is

a two-to-one map, which means that two distinct representative points in Ω are mapped onto the same point. In physical space this may occur when two distinct trajectories of disks are mapped onto the same trajectory, which is the case for the traditional Maxwell's demon (see section 8).

A two-to-one evolution map is irreversible as well as compressing phase space volume, but irreversibility is not necessary for compressing phase space volume. The next section shows that the membrane and disks system compresses phase space volume even though its dynamics are time-reversible.

Incompressibility of dynamics is a property of systems that have a proper Hamiltonian function (closed and isolated systems). Such systems conserve phase space volume according to Liouville's theorem [3]. A system that is the limit of Hamiltonian systems also conserves phase space volume. For example, a system of hard disks in a box (no membrane, only disks) conserves phase space volume [4], and it can be viewed as the limit of Hamiltonian systems with inverse power law potentials. By contrast, the membrane of equation 1 can not be the limit of Hamiltonian systems since it compresses and expands phase space volume, as is shown in the next section.

6 Phase Space Volume

The membrane and disks system is not ergodic even though its dynamics are chaotic, time-reversible, and energy conserving. The chaotic character results from the binary collisions. Time-reversibility results from the binary collisions and the membrane interaction of equation 1. One can easily verify that if the velocity of a disk is reversed after it has interacted with the membrane, the disk retraces its trajectory in all cases. The system of disks would be ergodic if the membrane interaction did not compress and expand phase space volume when a disk penetrates the membrane.

To examine the compressibility of dynamics, we can consider a membrane system that contains only one disk for simplicity. In other words, we have the same container and membrane as shown in figure 2 and we have a single disk bouncing around. The phase space Ω is three dimensional X, Y, Θ where X,Y is the position of the disk and Θ is the velocity angle. We examine the compressibility of dynamics using this one disk system.

Referring to figure 3 we see that the transformation of the velocity angle of an incident disk is a linear map that has unity slope (minus one). A unity slope suggests that the membrane map conserves volume in phase space. It turns out however that this is not correct. In order to calculate the compressibility of dynamics we have to consider the time evolution of all three X, Y, Θ

dimensions of the phase space together, and not only the Θ dimension that is shown in figure 3.

For concreteness we look at the time evolution of an infinitesimal phase space volume $\omega(X_1, Y_1, \Theta_1)$ centered at X_1, Y_1, Θ_1. We assume that the membrane is located at $X = 0$, and that the volume $\omega(X_1, Y_1, \Theta_1)$ is near the membrane with $X_1 > 0$, and that Θ_1 is inside the interval $3\pi/4 < \Theta_1 < \pi$. We assume that after a time interval of 1.0 (in appropriate units) every point X, Y, Θ in the volume $\omega(X_1, Y_1, \Theta_1)$ has penetrated the membrane from the right and moved to the left of the membrane. We denote by X', Y', Θ' the image of X, Y, Θ under the evolution map, and we have the following equation,

$$X' = X + \cos\Theta\, \Delta t_c - \sin\Theta\, (1 - \Delta t_c) \tag{7}$$

where Δt_c is the time it takes for point X, Y, Θ to move to the membrane and is equal to $-X/\cos\Theta$. Therefore,

$$\begin{aligned} X' &= -\sin\Theta - X\sin\Theta/\cos\Theta \\ Y' &= Y - X(1 + \sin\Theta/\cos\Theta) - \cos\Theta \\ \Theta' &= 3\pi/2 - \Theta \end{aligned} \tag{8}$$

To check whether the evolution map compresses the volume $\omega(X_1, Y_1, \Theta_1)$

we evaluate the Jacobian determinant of the above equations. We find,

$$\begin{vmatrix} -\sin\Theta/\cos\Theta & 0 & -\cos\Theta - X/\cos^2\Theta \\ -1-\sin\Theta/\cos\Theta & 1 & \sin\Theta - X/\cos^2\Theta \\ 0 & 0 & -1 \end{vmatrix} = \sin\Theta/\cos\Theta = \tan\Theta \tag{9}$$

For points X, Y, Θ inside the volume $\omega(X_1, Y_1, \Theta_1)$, the angle Θ is inside the interval $3\pi/4 < \Theta < \pi$. Thus, the Jacobian determinant is always less than one. In other words, the evolution map compresses phase space volume when a disk penetrates the membrane from right to left.

Figure 6 shows geometrically how the compression of phase space volume occurs when a disk penetrates the membrane. The figure is confined to two dimensions for practical reasons. The two rectangles shown in solid lines correspond to phase space volumes that are mapped onto each other under the evolution map — they correspond to $\omega(X, Y, \Theta)$ with the angle Θ chosen constant for all points. The vertical line $X = 0$ of figure 6 corresponds to the membrane. It is easy to see that the edges of the two rectangles (the original rectangle and its image under the evolution map) are equal between the two rectangles, but the angles of the two rectangles are not equal, and hence the areas of the two rectangles are not equal. Therefore phase space volume is compressed when penetrating the membrane from right to left and expanded when penetrating from left to right.

The significance of compressing phase space volume by the membrane map of equation 1 can be appreciated if we attempt to find a new membrane map that would result in a density difference while preserving phase space volume. It turns out that this is not possible, at least for piecewise differentiable maps. To see this, we seek a map $f(\Theta)$ mapping the impact velocity angle Θ to a new velocity angle $f(\Theta)$ so that the condition of incompressibility (Jacobian determinant unity) is satisfied. Repeating the above steps, equations 8 and 9 using $f(\Theta)$, we find

$$f(\Theta) = \sin^{-1}(\pm \sin\Theta \ + \ C) \ . \tag{10}$$

If C is zero, $f(\Theta)$ corresponds to a transparent membrane (i.e. no membrane at all) or a completely impenetrable membrane (i.e. a wall). If C is non-zero, then we get a membrane that maps velocity angles in a non-linear fashion. In analogy with the membrane of equation 1, we can apply the non-linear map $f(\Theta)$ to a selected region of velocity angles and block the remaining angles. In this way we hope that the probabilities of penetrating the membrane from left and right will be different from each other (see section 2). However, a simple calculation shows that the probabilities of penetrating the membrane from left and right must be equal to each other for all possible choices of the constant C in equation 10. For example let us suppose that $(\pi, \pi - d)$ is a region of velocity angles penetrating the membrane from the right side,

and $(f(\pi - d), f(\pi))$ is the image of this region under the membrane map $f(\Theta) = \pi - \sin^{-1}(\sin\Theta + C)$ for some positive constant $C <= 1 - \sin d$. Also let us assume that the remaining velocity angles are blocked and do not penetrate the membrane. Then the probability of an individual molecule to penetrate the membrane from the right (see section 3) is,

$$\int_{\pi-d}^{\pi} |\cos\theta| d\theta \ , \tag{11}$$

and the probability to penetrate from the left is,

$$\int_{f(\pi-d)}^{f(\pi)} |\cos\theta| d\theta \ . \tag{12}$$

An elementary integration gives

$$\begin{aligned} \text{Eq11} &= \sin(\pi - d) = \sin d \ , \\ \text{Eq12} &= \sin f(\pi - d) - \sin f(\pi) = \\ &\quad \sin[\sin^{-1}[\sin(\pi - d) + C]] - \sin[\sin^{-1}[\sin\pi + C]] = \\ &\quad \sin d \ . \end{aligned} \tag{13}$$

Hence, the probabilities of penetrating the membrane from the right and from the left are equal to each other. Therefore a membrane map (piecewise differentiable) that obeys incompressible dynamics can not create a density difference.

7 Microscopic Rectifiers

Although the membrane and disks system creates a density difference between two chambers initialized to have equal density, the second law of thermodynamics is not in danger. The microscopic dynamics of the membrane system compress phase space volume, and we can assume that the laws of nature prohibit the compression of phase space volume in an isolated system of disks. In this way the membrane and disks system relates the second law of thermodynamics to the incompressibility of microscopic dynamics. It shows that compressibility of dynamics can reduce the entropy of a statistical system if the system is designed to take advantage of the compression of phase space volume.

The significance of compressible dynamics can be appreciated further if we compare the membrane and disks system to a class of mechanisms known as microscopic rectifiers [8, 1]. These mechanisms are designed to extract work from the thermal motion of gas molecules by rectifying spontaneous variations in density between microdomains of gas. Microscopic rectifiers do not succeed as explained in references [8, 1] because the rectifying mechanism becomes thermalized and starts moving randomly in every possible way. In order to succeed the rectifying mechanism must be kept at a lower temperature than the surrounding gas molecules.

Interestingly the prevention of thermalization by a cooling process compresses phase space volume. To see how compression of phase space volume occurs when a microscopic rectifier is cooled, we examine the trapdoor system of reference [1]. For simplicity we consider a trapdoor system that contains only one disk. We denote by x, y, u, v the coordinates and velocity of the disk, and by X, U the position and velocity of the trapdoor. We assume that the cooling operation of reference [1] occurs at time T_c for $0 < T_c < 1$ (in appropriate units of time), and we consider the evolution of the system between times 0 and 1. The new state is given by the following equations,

$$\begin{aligned} x' &= x + uT_c + \eta u(1 - T_c) \\ y' &= y + vT_c + \eta v(1 - T_c) \\ u' &= \eta u \\ v' &= \eta v \\ X' &= X + UT_c + \epsilon U(1 - T_c) \\ U' &= \epsilon U \end{aligned} \tag{14}$$

where $0 < \epsilon < 1$ is the cooling parameter (a fixed number), and η is chosen so that the total energy of the system is conserved. If m, M are the masses of the disk and the trapdoor, η is given by the formula

$$\eta = \sqrt{1 + \frac{MU^2}{m(u^2 + v^2)}(1 - \epsilon^2)} \quad . \tag{15}$$

After some algebra we find that the Jacobian determinant of the evolution

equation 14 is equal to ϵ. In other words, phase space volume is compressed at a rate of ϵ. A similar calculation for a system containing a large number of disks, for example n disks, gives the Jacobian determinant

$$\text{Jacobian} = \epsilon \left(1 + (1-\epsilon^2)\frac{MU^2}{\sum_{i=1}^{n} m(u_i^2 + v_i^2)}\right)^{(n-1)} , \qquad (16)$$

which reduces to $\epsilon\,(1 + R(1-\epsilon^2))$ if we assume that the total energy of the disks is much larger than the energy of the trapdoor, and that R is the ratio of the energy of the trapdoor to the average energy of the disks. Further, if the trapdoor is colder than the disks, R is a small number and the Jacobian determinant is approximately equal to ϵ. Hence, cooling the trapdoor is accompanied by compression of phase space volume.

The difference between a cooled rectifier and the membrane and disks system is that the former is an open system while the latter is a closed system. In the open system compression of phase space volume is physically possible because it comes from external forces. The cooled rectifier is simply a heat engine. In the closed system we must assume that compression of phase space volume is not physically possible. Also if the open system is enclosed in a larger system that is closed and isolated (for example if a hot and a cold reservoir are used to perform the cooling), then the extended phase space must evolve incompressibly, and heat can only be converted to work while the system is approaching equilibrium.

Another difference between a cooled rectifier and the membrane and disks system is that the former is microscopically irreversible, while the latter is microscopically time-reversible. However, both systems are macroscopically irreversible.

In conclusion compressibility of dynamics brings together many different kinds of systems: heat engines, microscopic rectifiers, the membrane and disks system, and the traditional Maxwell's demon which also compresses phase space volume as is shown in the next section.

8 Maxwell's Demon

The present section discusses the relationship between the membrane system of equation 1 and the traditional Maxwell's demon. Maxwell's demon is an imaginary being (or device) that operates a microscopic door between two chambers containing disks [5, 1, 6, 7]. In its simplest version the demon opens the door when a disk is coming from the right, and closes the door when a disk is coming from the left. The demon's operations lead to a density difference between the two chambers, and the density difference can be used to extract work from the thermal motion of the disks. If the demon could operate in a closed cycle dissipating less energy than the work that can be extracted after the demon has finished its operations, then the second law

of thermodynamics would be violated. This conundrum has inspired a large volume of literature aimed at exorcising Maxwell's demon [6].

The most recent and most popular way of exorcising Maxwell's demon is the information theoretic approach which assumes that the demon must erase information in order to operate its trapdoor, and the erasure of information must be accompanied by a minimum amount of entropy production (energy dissipation). This simple assumption has led to many interesting ideas and conjectures regarding the role of information in physics, the relation between physical and algorithmic entropy, and the possibility of reversible dissipationless computation [9, 10, 6]. According to the information theoretic assumption Maxwell's demon is in accordance with the second law of thermodynamics because the reduction of entropy achieved by the demon is counterbalanced by an equal amount of entropy production that is necessary to implement the demon's irreversible operations.

The dynamics of the traditional Maxwell's demon are irreversible and also phase space volume compressing. This is because the traditional demon maps two distinct disk trajectories onto the same trajectory every time it interacts with a disk, by opening and closing its trapdoor. After the demon has interacted with a disk, it is impossible to distinguish whether the disk came from the opposite chamber or whether the disk bounced off the demon's

door.

In contrast to the traditional Maxwell's demon, the membrane and disks system is time-reversible. To view the membrane system as a Maxwell's demon, we imagine that the membrane interaction is the result of a *demon playing tennis* with the disks. The demon moves a tiny racket up or down so as to intercept the disk at the membrane line. In addition, the demon orients its racket in one of three possible orientations $+45$, -45, and $+90$ degrees, so as to reflect the incoming disk according to the map of equation 1 (also see figure 2). The mechanism used by the demon to position its racket is not specified because no mechanism exists according to the second law, and because there is no need to specify a mechanism in order to exorcise the demon. To exorcise the tennis demon as well as the traditional Maxwell's demon, we assume that compression of phase space volume is not physically possible in an isolated system of disks.

The novelty of the tennis demon is that its dynamics are time-reversible as opposed to the traditional Maxwell's demon whose dynamics are irreversible. The tennis demon shows that irreversibility of dynamics is not necessary for decreasing the entropy of an isolated system of disks. Compression of phase space volume is more general than irreversibility in the sense that irreversibility necessitates compression of phase space volume, but the reverse is not

true in continuum phase space. If the phase space were discrete, however, (discrete position velocity), then compression of phase space volume would necessitate irreversibility, and the tennis demon would become irreversible just like the traditional Maxwell's demon.

There is also another way of looking at the tennis demon which is in terms of an information theoretic model. Instead of viewing the demon as abstract dynamics, we view the demon as an agent that measures the position and velocity of an incoming disk, or somehow has this information, and acts according to this information in order to position the racket. By viewing the demon this way, the demon is no longer abstract dynamics.

In particular the information model separates the demon from the microscopic dynamics and gives the demon a macroscopic quality: the act of positioning a racket according to information. This property does not belong to the level of microscopic dynamics. For example, two disks do not measure or have information about each other in order to perform a collision. Instead the disks follow the microscopic laws of nature that are executed infinitely precisely and time-reversibly. By contrast, the model of the demon acting according to information introduces the macroscopic qualities of finite precision and irreversibility.

Viewing the tennis demon as an agent that acts according to information

means that the tennis demon can only position its racket with a precision that is limited by the amount of information that the demon has. To position its racket, the demon must know exactly where an incoming disk is going to intersect the membrane line. This corresponds to infinite information because the point where a disk intersects the membrane line belongs to a continuum. Thus the tennis demon can only be time-reversible if it operates on infinite information. The information model of the demon and the continuum dynamics of the demon inevitably lead to the need for infinite information.

The novelty of the tennis demon is that the infinite information needed by the demon is always available in the system, before and after the demon has struck an incoming disk with its racket. The infinite information is the state (continuum position velocity) of the incoming disk, and it evolves in a time-reversible fashion. If the demon can have this infinite information, then it is a reversible demon. If it can not have this infinite information (possibly because of quantum uncertainty), then it is an irreversible demon.

To label the tennis demon irreversible, however, is somewhat misleading. First, this terminology does not capture the nature of the tennis demon which is the reversible compression and expansion of phase space volume. Second, the demon's need for infinite information is partly the result of our

viewing the demon as an agent that acts according to information. The information model of the demon is not necessary for exorcising the demon. It is only necessary for investigating other questions such as measurement theory. Leaving such researches aside, the tennis demon can be understood and exorcised as a time-reversible Maxwell's demon at the level of abstract dynamics. The exorcism follows from the assumption that the laws of nature prohibit compression of phase space volume in an isolated system of disks.

References

[1] P.A. Skordos and W.H. Zurek, "Maxwell's demon, rectifiers, and the second law: Computer simulation of Smoluchowski's trapdoor", *Am. J. Phys.* October 1992.

[2] H.R. Pitt, *Measure and Integration for Use*, (Clarendon Press, New York 1985), pp. 63—69.

[3] K. Huang, *Statistical Mechanics*, (Second Edition, John Wiley Sons., New York, 1987), pp. 64—65.

[4] C. Cercignani, *The Boltzmann Equation and its Applications*, (New York, Springer-Verlag, 1988), pp. 13—19.

[5] W. Ehrenberg, "Maxwell's Demon", *Sci. Am.* **217** 103—110, (November 1967).

[6] H.S. Leff and A.F. Rex, "Resource Letter MD-1: Maxwell's demon", *Am. J. Phys.* **58** (3) 201—209, (March 1990) ; *Maxwell's Demon: Entropy, Information*, edited by H.S. Leff and A.F. Rex, (Hilger/Princeton U.P., Princeton 1990).

[7] J.C. Maxwell, *Theory of Heat*, (Longmans, Green, London 1871), pp. 308—309.

[8] R.P. Feynman, R.B. Leighton, and M. Sands, *The Feynman Lectures on physics - Vol.1*, (Addison-Wesley Publication, Reading, Massachusetts 1963, 1965), chapter I-46.

[9] C.H. Bennett, "Demons, Engines, and the Second Law", *Sci. Am.* **257**, 108—116, (November 1987).

[10] L. Szilard, "Uber die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen," Z. Phys. **53** 840—856 (1929); "On the Decrease of Entropy in a Thermodynamic System by the Inteven-tion of Intelligent Beings," translation in *Quantum Theory and Measurement*, edited by J.A. Wheeler and W.H.Zurek, (Princeton University Press, Princeton 1983), pp. 539—548.

Figure 1: A system of membrane and disks can be viewed as a time-reversible Maxwell's demon. The membrane, displayed as a dashed line, interacts with the disks according to a time-reversible and energy conserving rule which creates a density difference between the two chambers.

Figure 2: The membrane interaction rule is illustrated graphically. The vertical solid line denotes the membrane, and the other solid lines denote a division of the plane into octants. The dashed lines correspond to disk trajectories that penetrate the membrane.

Figure 3: The membrane map is shown as a map of the velocity angle. The impact velocity angle Θ is mapped to the new velocity angle Θ'. Both angles range from $-\pi$ to π. The dashed line segments correspond to trajectories that penetrate the membrane, while the solid line segments correspond to trajectories that do not penetrate the membrane.

Figure 4: The cumulative time average of the number of disks in the left chamber builds up from an initial value of 0.5 to approximately 0.7 (smooth curve). The fluctuating curve displays a sequence of short time averages, each one taken over $1.25 \times 10^{-10} sec.$

Figure 5: The cumulative time average of the number of disks in the left chamber is shown over a much longer time scale than Figure 3. The running averages (fluctuating curve) are taken over time intervals of $25 \times 10^{-10} sec.$

Figure 6: The membrane map compresses phase space volume. A rectangular region of points having identical velocity (impact angle) is compressed when the points penetrate the membrane from right to left.

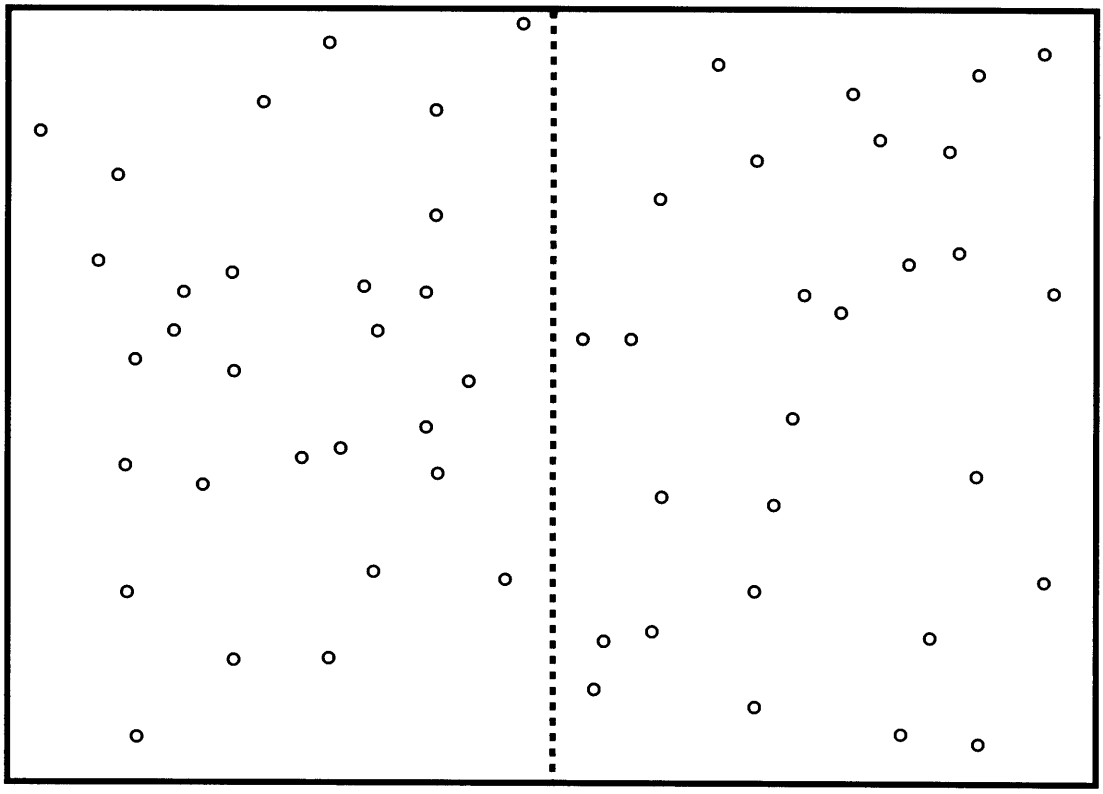

Fig. 1

skordos

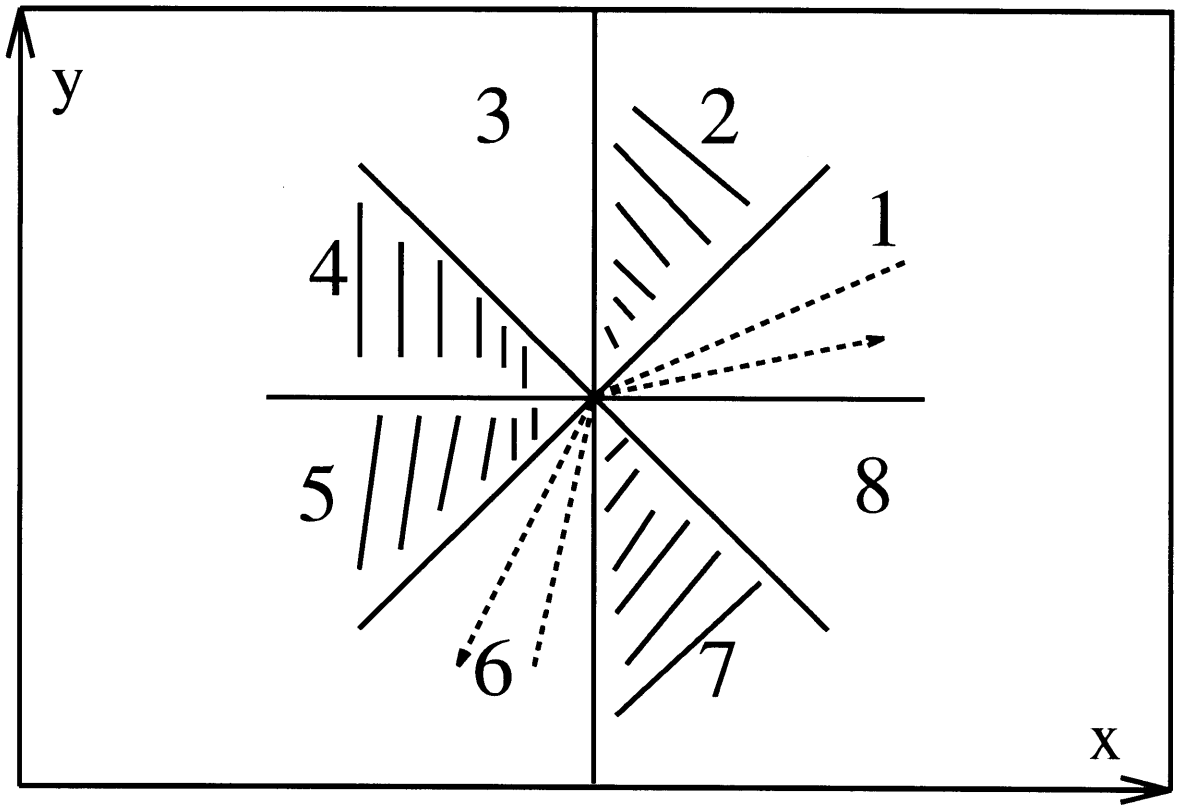

Fig. 2 skordos

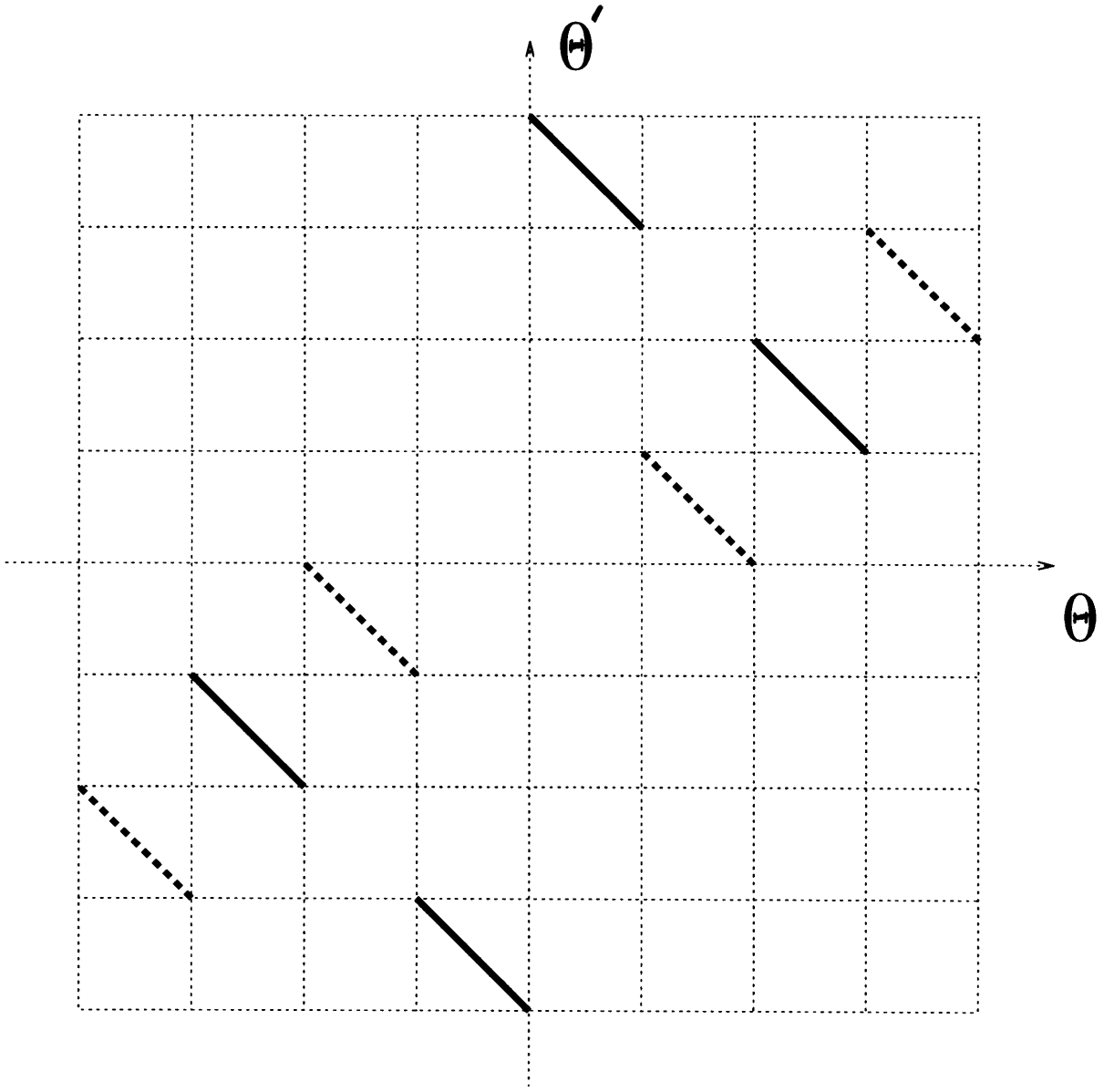

Fig. 3 skordos

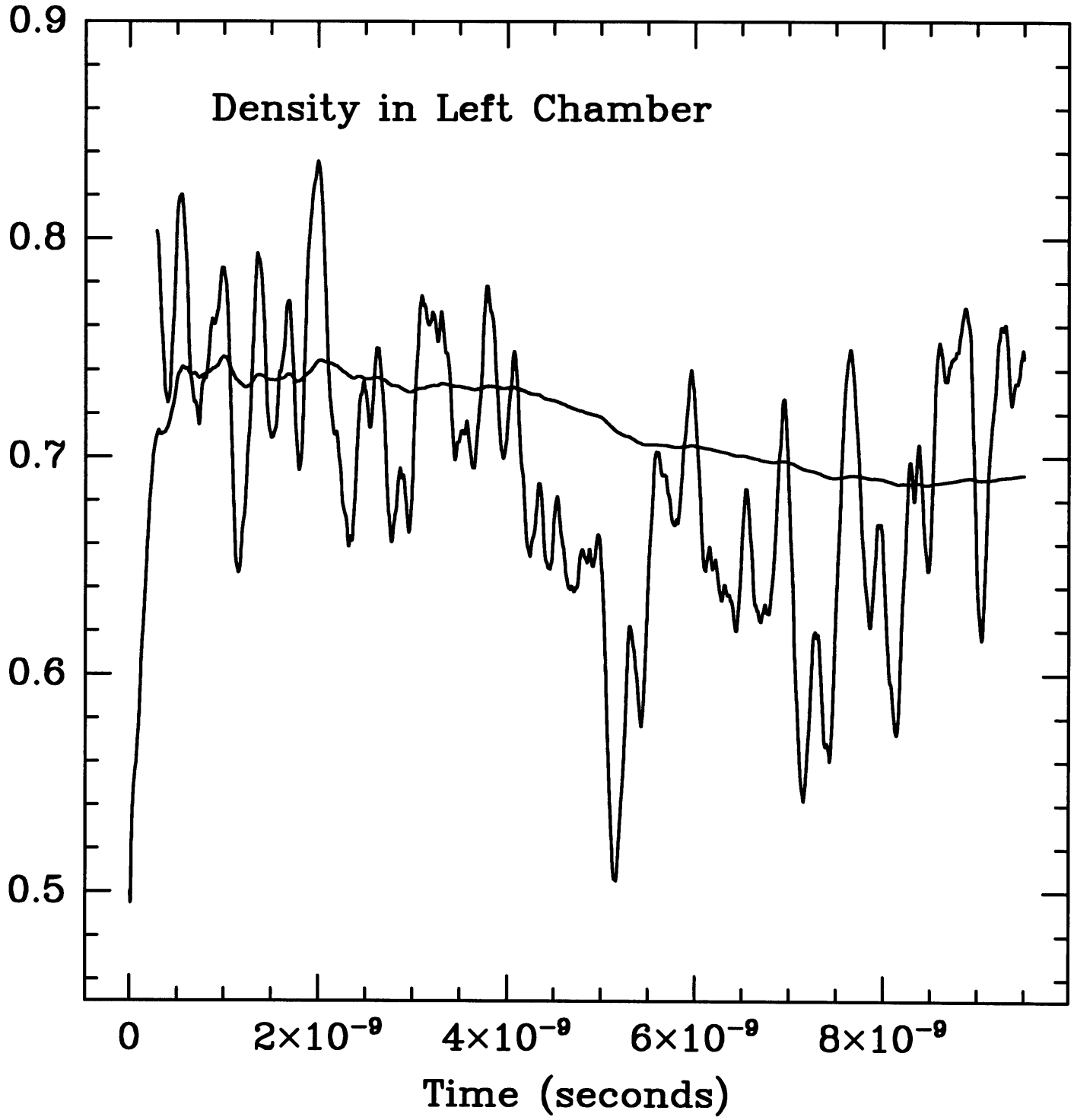

Fig. 4 skordos

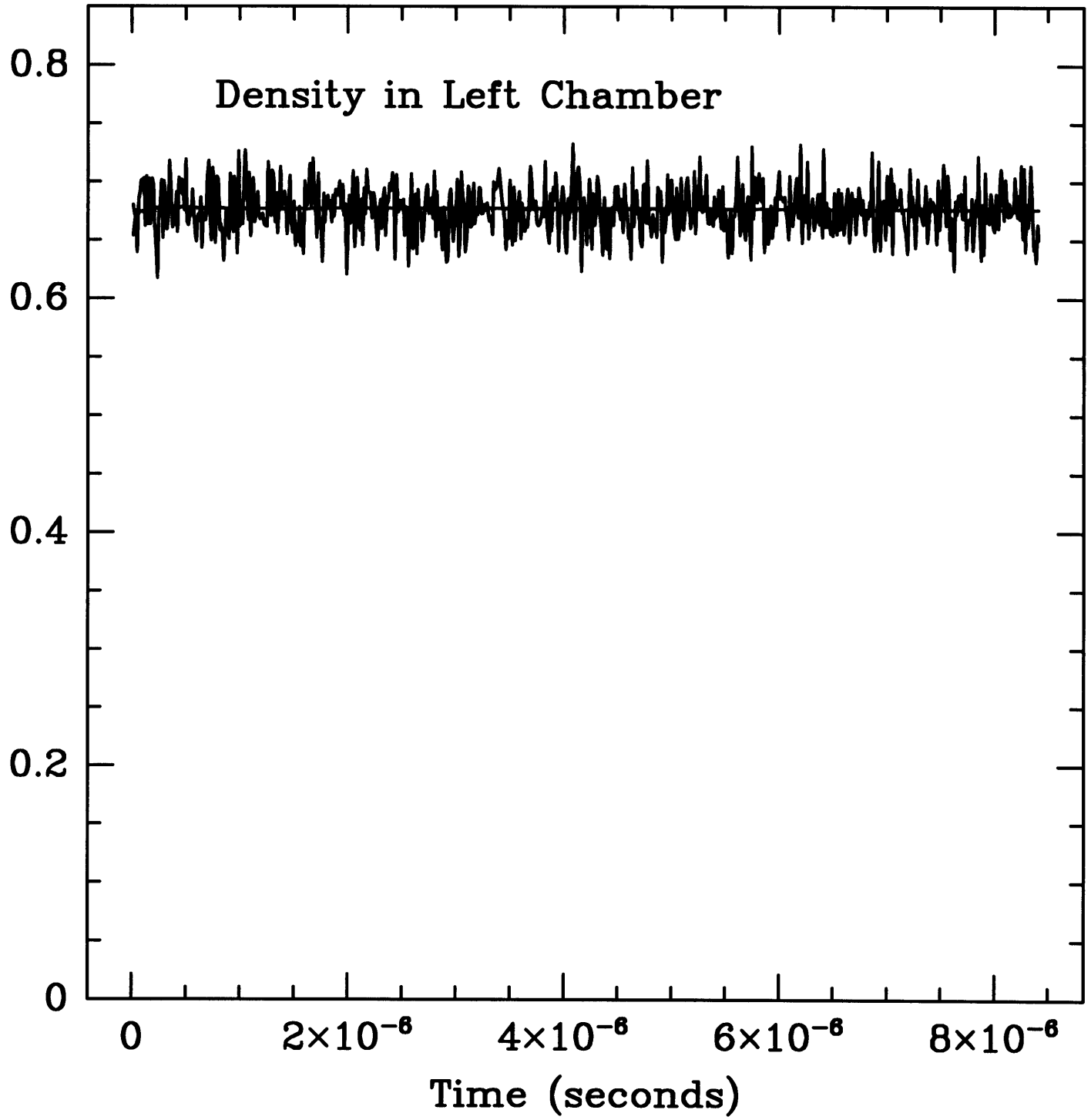

Fig. 5 skordos