Explaining irreversibility

Robert H. Swendsen[a)]

Department of Physics, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213

I describe an approach that helps students understand why the behavior of macroscopic systems distinguishes between past and future, even though the underlying classical or quantum microscopic equations are time-reversal invariant. This approach provides a qualitative understanding of irreversibility by an explicit calculation of irreversible behavior for a simple, time-reversal-invariant model system. The mathematics involved is accessible to upper-level undergraduate students.
© *2008 American Association of Physics Teachers.*

I. INTRODUCTION

When I was a student, I looked in vain for a clear explanation of how irreversible behavior in the real world could be compatible with the time-reversal-invariant equations of motion I was taught in classical and quantum mechanics. The literature I was referred to by my professors was mathematically intimidating, obscure, and not very illuminating.

This paper is an effort to help today's students who have similar questions. It proposes straightforward answers to the questions I had as a student, which might not be the same questions addressed by those who study the difficult mathematical problems associated with the field.

To begin with, the meaning of the term "irreversibility" is not as simple as it might seem. Much of the confusion and disagreement that has been associated with discussions of irreversibility comes from conflicting ideas of what the word means and what needs to be explained.[1] The irreversible behavior I wanted to understand, and which I want to explain to students, is what is observed. Every day we see that time runs in only one direction in the real world. If I drop my keys, they fall to the floor and stay there; keys lying on the floor do not suddenly jump into my hand. This asymmetry of time has seemed to many to be incompatible with the time-reversal invariance of the fundamental equations of classical and quantum physics.

The explanation of irreversibility given in this paper is based on the fact that macroscopic systems contain more than 10^{20} atoms, and that we have very limited information about the microscopic state. We might know that the atoms are in some sort of container, which restricts their positions, and we might know something about the temperature or total energy, which restricts their momenta. We might also have made observations of the uniformity of the density. It is unlikely that we have made more than a thousand measurements, and certainly less than a million. Therefore, the only reasonable description of our knowledge of the microscopic state of a macroscopic system is by probabilities.

The approach I have taken uses a simple thought experiment for which the microscopic equations of motion are time-reversal invariant, but the macroscopic behavior is nevertheless irreversible. The free expansion of a classical ideal gas satisfies these criteria. Many students — and some professors — are surprised that the ideal gas can exhibit irreversible behavior. After all, it has a number of constants of the motion equal to the number of components of the momentum. This surprise has the advantage of getting the students' attention.

Using a classical model removes any confusion between macroscopic irreversibility and the effects of quantum measurement, which might also show an asymmetry in time. It also simplifies the analysis. The mathematical solution for the properties of the model is straightforward, using only Fourier series expansions and Gaussian integrals.[2,3] Upper-level undergraduate students can do all the necessary calculations explicitly. Because the argument is transparent and nothing is assumed, students have found it satisfying and convincing.

The rest of the paper is organized in much the same order that I use in teaching this material. The next three sections will briefly review the history of the debate, discuss a trivial form of irreversibility that might otherwise cause confusion, and specify what I mean by the observation of equilibrium behavior. The exact solution, due to Frisch,[2] of the free expansion experiment is then given, followed by two sections showing how Zermelo's and Loschmidt's arguments are resolved. The properties of the solution to the free expansion experiment are then shown to suggest a generalization of the statistical mechanical definition of equilibrium to bring it in line with observations. The entropy is discussed, including a way to construct a time-dependent entropy that obeys the second law of thermodynamics. The paper ends with some concluding comments.

II. BOLTZMANN, LOSCHMIDT, AND ZERMELO

The history of the debate on irreversibility has been convoluted with responses to Boltzmann's famous H theorem.[4] Although Loschmidt[5] and Zermelo[6,7] aimed their objections at Boltzmann's kinetic equation rather than the apparent paradox itself, their objections represent aspects of the problem that must be addressed to give the students a complete picture.

Boltzmann pioneered the idea of using probability theory to understand the properties of materials. This use of probability theory was a very new idea in the 19th century, when physics — including thermodynamics — was regarded as describing phenomena that had no fundamental uncertainty. That Boltzmann connected the use of probability theory with the highly controversial idea of the existence of atoms made it even more unpalatable to many leading scientists of the day. Even Max Planck was initially an opponent of atomic theory, changing his views only in 1900 with his work on black body radiation.[8]

Boltzmann derived his equation for the time derivative of the distribution of atoms in the six-dimensional space of po-

sition and momentum by approximating the number of collisions between atoms ("Stosszahlansatz").[4] This approximation had the effect of turning the macroscopic dynamics into a Markov process, for which the development of the system depended only on the current state, and not on its history. Because Markov processes show irreversible behavior, Boltzmann's equation did also. In particular, Boltzmann showed that a particular quantity which he called H could not increase with time.

The first objection came from Boltzmann's friend, Johann Josef Loschmidt.[5] He pointed out that if the momenta of all the particles in an isolated system were to be instantaneously reversed, the system must retrace its trajectory. In particular, if Boltzmann's H function had been decreasing for a particular microscopic state, the value of H for the time-reversed system must increase, invalidating Boltzmann's conclusions. Loschmidt's argument is usually referred to by the German term "Umkehreinwand."[9]

Although Loschmidt's argument was aimed at Boltzmann's H theorem, it is very close to the central paradox. If every microscopic state on a trajectory that approaches equilibrium corresponds to a time-reversed state which is on a trajectory that moves away from equilibrium, shouldn't the two states be equally probable? The resolution of this apparent contradiction will be clear from the example that we will give.

Ernst Zermelo, a prominent mathematician who was not on very good terms with Boltzmann, raised a different objection more than 20 years after Boltzmann's original paper.[6,7] His argument was based on a recently derived theorem by Poincaré which required the behavior of any isolated classical system to be quasi-periodic.[10] That is, the system must return repeatedly to points that are arbitrarily close to its original configuration. Such behavior is incompatible with Boltzmann's H theorem, which predicts that the system will never leave equilibrium. Zermelo's argument is referred to by the German term "Wiederkehreinwand."[11]

Zermelo's Wiederkehreinwand is sometimes dismissed with a comment about the recurrence time being far too long to worry about. Some might find this argument sufficient, but we will see that we can deal with Zermelo's objection more directly on the basis of the example of free expansion.

III. A TRIVIAL FORM OF IRREVERSIBILITY

There is a danger of confusing the irreversibility observed in daily life with a trivial form of irreversibility that appears only in infinite systems. Consider a particle moving in empty space. At some time it is observed to be located in some finite region. Because the particle is moving, it will eventually leave this region. If space is infinite, it will not return. Therefore, we have found irreversible behavior, but of a trivial sort. This trivial form of irreversibility is also present in the propagation of sound waves away from a source in an infinite harmonic crystal or the radiation of light from a star. This irreversibility is quite general in infinite systems, whether classical or quantum. An open system also displays this trivial form of irreversibility, because it is really just a part of an infinite system.

To avoid confusing trivial and non-trivial irreversibility, it is important to restrict the discussion to an isolated, finite system.

IV. EQUILIBRIUM AND FLUCTUATIONS

As mentioned in Sec. I, the only kind of irreversibility I will discuss is that which is observed. Therefore, the concept of equilibrium must also be defined by observation. As an example, we might say that a glass of water has come to equilibrium when there is no measurable deformation of the water's surface. Naturally, our measurements would have limited resolution. We might be able to detect ripples with an amplitude less than a millimeter, but not less than a micron. Measurements of the wavelengths of ripples would be similarly limited.

In general, we will measure the uniformity of the system with a spatial resolution determined by our experimental equipment. Each measurement would also have an uncertainty determined by our equipment. When these imperfect measurements yield unchanging values, we would say that the system has come to equilibrium. No other type of irreversibility has ever been observed. Consequently, no other type of irreversibility requires explanation.

It is important to keep in mind that the properties of a system in equilibrium will fluctuate. The fluctuations of a macroscopic system will almost always be far smaller than the uncertainty of any measurement. However, there is always an extremely small, but non-zero probability of a very large fluctuation.

It is interesting to speculate on the reaction of an observer to the occurrence of a huge fluctuation, such as a set of keys, once dropped, jumping back into someone's hand. Because it would be unreasonable to hope that the phenomenon might be reproducible, it would certainly be regarded as either an experimental error or a miracle, depending on personal beliefs.

V. FREE EXPANSION OF A CLASSICAL IDEAL GAS

The central example in this paper is the free expansion of a classical ideal gas. Naturally, the magnitudes of the momenta of the individual particles are constants of the motion and do not change. However, the time development of the local density exhibits irreversible behavior.

Consider a system containing N non-interacting particles in a total volume V. Assume that the walls of the box are perfectly reflecting and isolate the gas from the rest of the universe. In the initial state, the gas is confined to a smaller volume V_0 by an inner wall. At time $t=0$, the confining wall is removed. For convenience, consider a rectangular box and align the coordinate axes with the sides. Let the inner wall which confines the atoms to the subvolume V_0 for $t \leq 0$ be perpendicular to the x direction and located at $x=L_0$, and let the length of the box in the x direction be L. The dependence of the probability distribution on the y and z coordinates does not change with time, so we need to consider only the x dependence, and we are left with a one-dimensional problem.

At time $t=0$ the confining wall at L_0 is removed, and the particles are free to move in the interval $0 < x < L$. We now have the advantage of being able to use an exact solution for the time dependence of the probability distribution due to Frisch.[2]

Following Frisch,[2] we can map the problem onto a box of length $2L$ with periodic boundary conditions. The particles are then confined to the region $-L_0 < x < L_0$ for $t \leq 0$, and $-L < x < L$ for $t > 0$. When a particle bounces off a wall in
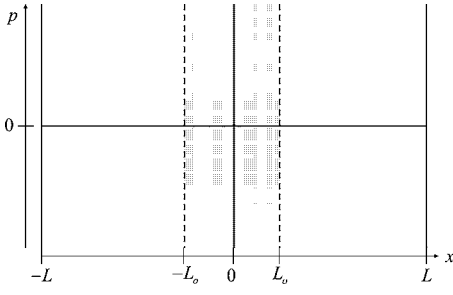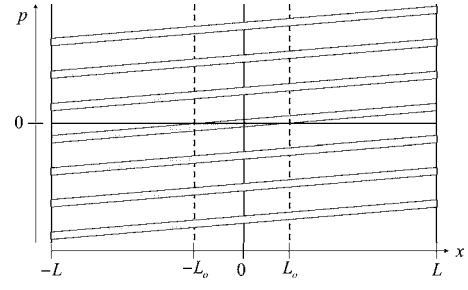
Fig. 1. The initial probability distribution at $t=0$.

Fig. 2. The probability distribution for $t>0$.

the original system, it corresponds to a particle passing between the positive and negative sides of the box without a change of momentum.

As mentioned in Sec. I, I will make the key assumption that we have very limited information about the initial microscopic state, so it must be described by a probability distribution. To be specific, assume that before the inner wall is removed, all initial positions and momenta are independent, and all particles have the same initial probability distribution. Let the spatial probability distribution be uniform, and the momenta have a Maxwell–Boltzmann distribution.

$$f_N(\{x_j,p_j|j=1\dots N\},t=0) = \prod_{j=1}^{N} f(x_j,p_j,t=0)$$
$$= \prod_{j=1}^{N} g_0(x_j)h_0(p_j), \qquad (1)$$

where

$$g_0(x_j) = \begin{cases} 1/L_0 & (-L_0 < x < L_0) \\ 0 & (|x| \ge L_0) \end{cases}. \qquad (2)$$

This initial probability distribution for the periodic system is shown in Fig. 1. Note that this probability distribution is time reversal invariant, so that there is no time asymmetry in the initial conditions.

Because the particle probabilities are independent, we can restrict our attention to the distribution function of a single particle. The periodic boundary conditions then allow us to make a Fourier series expansion of the initial conditions to obtain

$$g(x) = g_0 + \sum_{n=1}^{\infty} g_n \cos\left(\frac{\pi n}{L}x\right), \qquad (3)$$

where no sine terms enter because of the symmetry of the initial conditions.

We can find the coefficients g_n by the standard procedure of multiplying Eq. (3) by $\cos(\pi n'x/L)$ and integrating. The result is

$$\int_{-L}^{L} g(x)\cos\left(\frac{\pi n'}{L}x\right)dx = \int_{-L}^{L} g_0 \cos\left(\frac{\pi n'}{L}x\right)dx + \sum_{n=1}^{\infty} g_n$$
$$\times \int_{-L}^{L} \cos\left(\frac{\pi n}{L}x\right)\cos\left(\frac{\pi n'}{L}x\right)dx. \qquad (4)$$

For $n'=0$ Eq. (4) gives

$$g_0 = \frac{1}{L} \qquad (5)$$

and for $n' \ge 1$

$$g_{n'} = \frac{2}{n'\pi L_0} \sin\left(\frac{n'\pi}{L}L_0\right). \qquad (6)$$

The time development for a given momentum is now simple.

$$f(x,p,t) = f\left(x-\frac{pt}{m},p,0\right) \qquad (7a)$$

$$= h_0(p)\left[g_0 + \sum_{n=1}^{\infty} g_n \cos\left(\frac{n\pi}{L}\left(x-\frac{pt}{m}\right)\right)\right]. \qquad (7b)$$

Figure 2 shows the probability distribution at a time $t>0$. The shaded areas, representing non-zero probabilities, tilt to the right. As time goes on, the shaded areas become increasingly flatter and closer together.

We have assumed that the initial probability distribution for the momentum is given by the Maxwell–Boltzmann distribution.

$$h_0(p) = \left(\frac{\beta}{2\pi m}\right)^{1/2} \exp\left[-\beta\left(\frac{p^2}{2m}\right)\right], \qquad (8)$$

where T is the temperature, and $\beta=1/k_BT$. We substitute Eq. (8) in Eq. (7a) and integrate over the momenta to calculate the probability density as a function of position and time.

$$g(x,t) = \int_{-\infty}^{\infty} f(x,t)dp \qquad (9a)$$

$$= \int_{-\infty}^{\infty} h_0(p)L^{-1}dp + \sum_{n=1}^{\infty} g_n \int_{-\infty}^{\infty} h_0(p)$$
$$\times \cos\left(\frac{n\pi}{L}\left(x-\left(\frac{p}{m}\right)t\right)\right)dp. \qquad (9b)$$

The first integral is trivial because $h_0(p)$ is normalized.

$$\int_{-\infty}^{\infty} h_0(p)dp = 1. \qquad (10)$$

The integrals in the sum are a bit tedious, but not particularly difficult if the identity $\cos\theta = (e^{i\theta}+e^{-i\theta})/2$ is used to write them in the form

$$\int_{-\infty}^{\infty} \exp\left[-\beta\frac{p^2}{2m} \pm \frac{n\pi}{L}i\frac{p}{m}t\right]dp. \tag{11}$$

We complete the square in the exponent and obtain

$$g(x,t) = L^{-1} + \frac{2}{\pi L_0}\sum_{n=1}^{\infty}\frac{1}{n}\sin\left(\frac{n\pi}{L}L_0\right)\cos\left(\frac{n\pi}{L}x\right)$$
$$\times\exp[-\lambda_n^2 t^2], \tag{12}$$

where

$$\lambda_n^2 = \frac{n^2\pi^2}{2mL^2\beta} = \frac{n^2\pi^2}{2mL^2}k_B T. \tag{13}$$

Because the temperature is related to the one-dimensional root-mean-square velocity by

$$\frac{1}{2}k_B T = \frac{1}{2m}\langle p^2\rangle = \frac{m}{2}\langle v^2\rangle = \frac{m}{2}v_{\text{rms}}^2, \tag{14}$$

Eq. (13) can be written as

$$\lambda_n^2 = \frac{n^2\pi^2}{2L^2}v_{\text{rms}}^2. \tag{15}$$

Equations (12)–(15) can also be expressed in terms of the characteristic time for a particle traveling with the speed v_{rms} to cross the box.

$$\tau = \frac{L}{v_{\text{rms}}}. \tag{16}$$

In terms of τ the time-dependent local probability density becomes

$$g(x,t) = L^{-1}\left[1 + \frac{L}{L_0}\sum_{n=1}^{\infty}\left(\frac{2}{n\pi}\right)\sin\left(\frac{n\pi}{L}L_0\right)\cos\left(\frac{n\pi}{L}x\right)\right.$$
$$\left.\times\exp\left[-\left(\frac{\pi^2 n^2}{2}\right)\left(\frac{t^2}{\tau^2}\right)\right]\right]. \tag{17}$$

Because $|(2/n\pi)\sin(\frac{n\pi}{L}L_0)\cos(\frac{n\pi}{L}x)| \leq \frac{2}{n\pi}$ for all x, the sum converges rapidly for $t > 0$. At $t = \tau$, the nth term in the sum is less than $(2/n\pi)\exp[-\pi^2 n^2/2]$, and even the leading non-constant term has a relative size of less than $0.005L/L_0$. The rapidity of convergence is striking. The inclusion of interactions in the model would actually slow down the approach to equilibrium because it would involve diffusive processes with characteristic times that grow as L^2.

We can calculate the energy density as a function of position and time by including a factor of $p^2/2m$ in the integral over the momentum.

$$U(x,t) = \int_{-\infty}^{\infty} f(x,p,t)\frac{p^2}{2m}dp \tag{18a}$$

$$= \int_{-\infty}^{\infty} h_0(p)\frac{p^2}{2m}L^{-1}dp$$
$$+ \sum_{n=1}^{\infty} g_n \int_{-\infty}^{\infty} h_0(p)\frac{p^2}{2m}\cos\left(\frac{n\pi}{L}\left(x - \frac{pt}{m}\right)\right)dp. \tag{18b}$$

The integrals are more tedious than those in Eq. (9a), but not particularly difficult. We find

$$U(x,t) = \frac{1}{2L\beta} - \sum_{n=1}^{\infty}\left(\frac{n^2\pi^2 t^2 - L^2 m\beta}{L^2 L_0 mn\pi\beta^2}\right)$$
$$\times\sin\left(\frac{n\pi L_0}{L}\right)\cos\left(\frac{n\pi x}{L}\right)e^{-\pi^2 n^2 t^2/2\tau^2}. \tag{19}$$

After the internal wall is removed, the energy distribution is not uniform and not proportional to the particle density, because the faster particles move more rapidly into the region that was originally vacuum. However, the energy density converges rapidly to the expected constant, $1/2L\beta = k_B T/2L$, as $t \to \infty$.

VI. ZERMELO'S WIEDERKEHREINWAND REVISITED

Students usually find an explanation of Poincaré recurrences helpful in understanding the reconciliation of Zermelo's Wiederkehreinwand with the monotonic approach to equilibrium shown by Eqs. (17) and (19). It is easy to see how quasi-periodic behavior can arise in the ideal gas model.

First consider two particles with speeds v_1 and v_2. They will each return to their original states with periods $\tau_1 = 2L/v_1$ and $\tau_2 = 2L/v_2$, respectively. In general, the ratio τ_1/τ_2 will be irrational, but it can be approximated to arbitrary accuracy by a rational number, $\tau_1/\tau_2 \approx n_1/n_2$, where n_1 and n_2 are integers. Therefore, after a time $\tau_{1,2} = n_2\tau_1 \approx n_1\tau_2$, both particles will return to positions and velocities arbitrarily close to their initial states if sufficiently large integers n_1 and n_2 are used in approximating the ratio τ_1/τ_2.

Now add a third particle with speed v_3 and period $\tau_3 = 2L/v_3$. A rational approximation, $\tau_{1,2}/\tau_3 \approx n_{1,2}/n_3$, will give us an approximate recurrence after a time $\tau_{1,2,3} = n_3\tau_{1,2} \approx n_{1,2}\tau_3$. Because n_1, n_2 and n_3 will usually be large numbers, $\tau_{1,2,3}$ will usually be a long time. If we repeat this procedure for 10^{23} particles, we will usually arrive at a recurrence time that would be enormous even in comparison to the age of the universe.[12] It might be said that we are not interested in such extremely long times, but it is interesting that Eqs. (17) and (19) do not exhibit the quasi-periodic behavior that Poincaré recurrence might seem to require.

The resolution of the apparent contradiction lies in our lack of knowledge of the exact initial velocities and the extreme sensitivity of the Poincaré recurrence time to tiny changes in initial velocities. Even with far more detailed information than we would ever obtain in a real experiment, we would not be able to predict a Poincaré recurrence time with an uncertainty of less than many ages of the universe.

Because the solution for the free expansion of an ideal gas is exact, Poincaré recurrences are automatically included. Because we cannot predict them, they appear as extremely rare large fluctuations. Because such large fluctuations are always possible in equilibrium, there is no contradiction between the observed irreversible behavior and Poincaré recurrences.

VII. LOSCHMIDT'S UMKEHREINWAND REVISITED

Loschmidt's Umkehreinwand was directed against Boltzmann's equation, which was not time-reversal invariant. However, the exact solution to the free expansion of an ideal gas retains the time-reversal properties of the microscopic equations. If we reverse all velocities in the model some time after the inner wall is removed, the particles return to their original positions.

Under normal conditions, the reversal of molecular velocities is difficult. However, the reversal of spin precession, known as the spin-echo effect,[13] can be accomplished and is fundamental to magnetic resonance imaging. After an initial magnetic pulse aligns the spins in a sample, interactions between the spins and inhomogeneities in the sample cause the spins to precess at different rates and lose the coherent alignment. If no further action is taken, the spins will not return to a coherent aligned state. However, if a second magnetic pulse is used to rotate the spins by 180° at a time t after the first pulse, it effectively reverses the direction of precession of the spins. The spins realign (or "refocus") after a total time $2t$ after the initial pulse. The reversal of the time development of an "irreversible" process is therefore not a flaw in the free-expansion model, but a reflection of the reality that such experiments can be carried out in practice, at least to a limited extent.

VIII. WHAT IS EQUILIBRIUM?

An interesting feature of the exact solution to the free-expansion model can be seen in Fig. 2. Even though the shaded regions, indicating non-zero probabilities, become progressively thinner and closer together, the local probability density at any point along the trajectory of the system remains constant. This behavior of the probability density is consistent with Liouville's theorem which states that the total time derivative of the probability distribution vanishes.[14] This property of all isolated Hamiltonian systems has caused difficulties for those who would like to define equilibrium in terms of a smooth distribution in phase space. Although we assumed that our model experiment started with a smooth probability distribution given in Eqs. (1), (2), and (8), it will certainly never evolve to one. The distribution will retain the layered structure shown in Fig. 2 for arbitrarily long times and will never become smooth.

The difficulty is due to a confusion about the direction of inference. It has been amply demonstrated that a canonical distribution accurately describes an equilibrium state. However, that does not imply that a system in equilibrium can only be described by a canonical distribution. The probability distribution for our model at long times will give the same macroscopic predictions as the canonical distribution, even though the two distributions are different.

To clarify this point, it can be useful to consider two successive free expansion experiments. Begin with a gas con-fined to a box of length L_0 with a canonical distribution of the momenta. At $t=0$, carry out a free expansion to a box of length $L_1 > L_0$. At a later time, $t_1 > 0$, carry out a second free expansion to a box of length $L_2 > L_1$. At the beginning of the second free expansion ($t=t_1$), the probability distribution function will look like Fig. 2, and the slope of the stripes will be $m/t_1 > 0$. As the probability distribution function develops for $t > t_1$, the slope of these stripes will continue to decrease as m/t. However, they will be contained in an envelope of similar stripes with a slope $m/(t-t_1) > 0$. Both of these slopes will be positive and go to zero as $t \rightarrow \infty$.

The macroscopic behavior of the second free expansion will be essentially the same as that of a free expansion from a canonical distribution. At the microscopic level, however, the probability distribution will differ from both the canonical distribution and the distribution for a single free expansion. A series of free expansions can create a probability distribution with a complicated structure of stripes, but the macroscopic approach to equilibrium would not be affected.

Students should be released from the notion that thermal equilibrium corresponds to a unique distribution in phase space and taught to understand it in terms of time-independent macroscopic behavior. The free expansion of the ideal gas is an excellent way for them to see the fundamental issues.

IX. ENTROPY

Liouville's theorem has also caused difficulties for the traditional textbook definition of the entropy as the logarithm of a volume in phase space. The theorem requires that this volume remain constant, so the traditional expression for the entropy cannot increase in an isolated system. This result seems to violate the second law, but it is correct. The traditional definition of entropy is related to the total information we have about the system — not the thermodynamic information about the current and future behavior of the system. The information that the system was initially confined to a smaller volume is contained in the layered structure of the probability density shown in Fig. 2. Because that information does not change, the traditional entropy does not change. The apparent violation of the second law arises because the traditional entropy does not correspond to the thermodynamic entropy.

If we use Boltzmann's definition of the thermodynamic entropy in terms of the probability of a macroscopic state,[4,15] we obtain an expression that increases with time as expected for the thermodynamic entropy. The specific form of the entropy depends on what we are measuring in the experiment. Because we are interested in the time development of the system as it approaches a uniform state, it would be reasonable to observe its properties as a function of position. We divide the system into M subsystems, each with length $\Delta L = L/M$, and measure the energy and density in each subsystem. For convenience, assume that the first m subsystems are inside the inner wall before the free expansion began, so that $L_0 = mL/M$.

If M is large, we can assume that the energy and density are uniform across a subsystem. The number of particles in the jth subsystem, $N_j(t)$, is given by ΔL times the expression in Eq. (12), and the energy in the jth subsystem, $E_j(t)$, is given by ΔL times the expression in Eq. (19). If the subsystems are sufficiently large so that $1/\sqrt{N_j(t)}$ is much

smaller than the relative error in the experimental measurements, they can be regarded as macroscopic systems. Their individual entropies are then given by the equilibrium entropy of a one-dimensional, classical ideal gas.

$$S(E_j(t), L/M, N_j(t)) = N_j(t) k_B T \left[\ln\left(\frac{L/M}{N_j(t)}\right) \right.$$
$$\left. + \frac{1}{2} \ln\left(\frac{E_j(t)}{N_j(t)}\right) + X \right], \quad (20)$$

where X is a constant. The total time-dependent entropy of the whole system is then given by

$$S(\{E_j(t)\}, L, N_j(t), M) = \sum_{j=1}^{M} S(E_j(t), L/M, N_j(t)). \quad (21)$$

Equation (21) has the properties that at $t=0$, it takes on the value of the initial entropy before the inner wall was removed,

$$S(\{E_j(0)\}, L, N_j(0), M) = S(E, L_0, N), \quad (22)$$

and as $t \to \infty$, it goes rapidly to the equilibrium entropy of the full system.

$$\lim_{t \to \infty} S(\{E_j(t)\}, L, N_j(t), M) = S(E, L, N). \quad (23)$$

These two properties are independent of the number M of subsystems that are observed.

X. CONCLUSIONS

I have presented a way to explain how irreversible processes seen in the real world are consistent with the time-reversal-invariant equations that we teach. The mathematics of the argument is at a level that an upper-level student should be able to follow.

The explanation of irreversible phenomena is not complete. I have restricted myself to the question of the apparent conflict between microscopic and macroscopic laws, while ignoring the effects of interactions. My experience is that once students understand why there is no fundamental contradiction between microscopic time-reversal invariance and macroscopic irreversibility, they have no trouble understanding how interactions lead to diffusion, dissipative phenom-

ena, and the establishment of the Maxwell-Boltzmann distribution of momenta. Molecular dynamics simulations are particularly good for helping students see how equilibration takes place with interacting particles.

My hope is that this approach to teaching the origins of observed irreversibility will help take the mystery out of the subject. Students should be confident that the foundations of statistical mechanics are solid and reliable for explaining observed phenomena, including irreversible behavior.

[a] Electronic mail: swendsen@cmu.edu

[1] P. Coveney and R. Highfield, *The Arrow of Time* (W. H. Allan, London, 1990). This book contains a good discussion of the debate about irreversibility.

[2] H. L. Frisch, "An approach to equilibrium," Phys. Rev. **109**, 22–29 (1958).

[3] A. Georgallas, "Free expansion of an ideal gas into a box: An exactly solvable approach to equilibrium," Phys. Rev. A **35**, 3492–3498 (1987).

[4] L. Boltzmann, "Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen," Wien. Ber. **66**, 275–370 (1872). The German word "Stoss" refers to a collision, "Zahl" means number, and "Ansatz" has become an English word for the starting point, or initial approximation.

[5] J. Loschmidt, "Über den Zustand des Wärmegleichgewichtes eines Systems von Körpern mit Rücksicht auf die Schwerkraft," Sitzungsber. Akad. Wiss. Wien, Math.-Naturwiss. Kl., Abt. 2A **73**, 128–142 (1876).

[6] E. Zermelo, "Über einen Satz der Dynamik und die mechanische Wärmelehre," Ann. Phys. **57**, 485–494 (1896).

[7] E. Zermelo, "Über mechanische Erklärungen irreversibler Vorgänge. Eine Antwort auf Hrn. Boltzmann's 'Entgegnung'," Ann. Phys. **59**, 793–801 (1896).

[8] M. Planck, "Über das Gesetz der Energieverteilung im Normalspectrum," Ann. Phys. **309**, 553–563 (1901).

[9] "Umkehr" means "reversal," referring to the reversal of velocities in Loschmidt's argument. "Einwand" means "objection."

[10] H. Poincaré, "Sur le probleme des trois corps et les équations de la dynamique," Acta Math. **13**, 1–270 (1890).

[11] "Wiederkehr" means "return" or "recur," referring to the return to the original state in Zermelo's argument.

[12] If the period of each molecule divided into the longest period is an integer, the recurrence time will be equal to the longest period. It will not be astronomically long. Therefore, a probabilistic argument is necessary to establish that the recurrence time is extraordinarily long.

[13] E. L. Hahn, "Spin echos," Phys. Rev. **80**, 580–594 (1950).

[14] R. K. Pathria, *Statistical Mechanics* (Pergamon, New York, 1972), pp. 34–36.

[15] R. H. Swendsen, "Statistical mechanics of colloids and Boltzmann's definition of the entropy," Am. J. Phys. **74**, 187–190 (2006).