International Journal of Theoretical Physics, Vol. 16, No. 4 (1977), pp. 287-298

Thermodynamic Arrow for a Mixing System[1]

L. S. SCHULMAN and R. SHTOKHAMER

Physics Department, Indiana University, Bloomington, Indiana 47401, and Physics Department, Technion, Haifa, Israel

Received: 1 *February* 1976

Abstract

The purpose of this paper is to study the appearance of time asymmetry in dynamical systems. The systems are harmonic oscillators and a certain mixing flow on the torus. The asymmetry is a kind of frictional force, but we emphasize that the boundary conditions, a usual source of asymmetry in studies of this sort, are taken to be time symmetric. For the mixing flow the response of the system, as reflected in its entropy as a function of time, occurs only subsequent to the "friction," while for the oscillators the effects are both before and after. Some general discussion also takes up the question of which of the foregoing systems is a better model of the physical world for purposes of correlating arrows of time.

1. *Introduction*

In an earlier paper (Schulman, 1973) one of us examined a thought experiment of Gold (1962a, b) which was supposed to correlate the thermodynamic arrow of time with the cosmological arrow.[2] According to Schulman (1973), a correct framework for demonstrating such a correlation is the following: For some model dynamical system sufficient (but not excessive) data about the state of the system are given at two different times ($t = 0$ and $t = T_1 + T_2$), and the state is studied at intervening times. At some intermediate time T_1 a time reversal noninvariant perturbation is applied to the system. If the data at 0 and $T_1 + T_2$ are roughly the "same" (in what sense will become evident), then the only asymmetry (or arrow) in the problem is the perturbation at T_1. To demonstrate a correlation of this arrow with some other arrow, one would

[1] Work supported in part by U.S. Army Research Office (Durham, N.C.).

[2] If one agrees with E. R. Harrison (1974) it would be more accurate to say that Gold uses a "dark sky arrow" rather than an arrow determined by the expansion of the universe (Gold uses the fact that more photons leave than arrive). In any case, we continue to use the terminology "cosmological arrow."

then have to show the emergence of the other arrow in the behavior of the system. In particular, it would be interesting to see whether the response (by the system) to the perturbation occurs both before and after the perturbation or only on one side.[3]

In Schulman (1973) a collection of harmonic oscillators was studied and no arrow found. That is, boundary values for the oscillators were given at $t = \pm T_0$ and friction allowed to act at $t = 0$. By considering oscillators with many different frequencies a sort of equilibrium was reached for $-T_0 + \tau < t < -\tau$ and $\tau < t < T_0 - \tau$, where τ is some characteristic equilibration time. However, the system was found to depart from equilibrium both for $0 < t < \tau$ and $-\tau < t < 0$. Arguments were also given for this noncausal behavior to occur in dynamical systems more general than harmonic oscillators.

In this paper we study another model, but are able to report that this system does show causal behavior. The model is a certain flow on a torus (details below) and was selected because the flow is known to be mixing. It was felt that the mixing property would facilitate the appearance of phenomena associated with time asymmetry, and this indeed appears to be the case. A perturbation of a time asymmetric type is applied and the system found to be out of equilibrium only after the perturbation.

Section 4 contains some general arguments which explore the question of which model is a better description of natural phenomena, the oscillators or the flow on the torus, with the object of determining whether optimism or pessimism is in order for the program of correlating arrows.

2. *Dynamical Systems*

Two dynamical systems will be considered: harmonic oscillators and discrete flow on a torus.

A harmonic oscillator with natural frequency ω has Hamiltonian

$$H = \tfrac{1}{2}(p^2 + \omega^2 q^2) \tag{2.1}$$

where p and q are Cartesian coordinates in its phase space. In polar coordinates,

$$E = \tfrac{1}{2}(p^2 + \omega^2 q^2)$$

$$\theta = \tan^{-1}(\omega q/p) \tag{2.2}$$

[3] In Gold's thought experiment the system is a star confined to a box for millions or perhaps billions of years. At some stage a window is opened and photons escape. The expansion of the universe and darkness of the night sky are what cause this asymmetric behavior, namely, the net escape rather than the absorption of photons. The question for this system is whether it departs from equilibrium only after the opening of the window, or also before. See Schulman (1973) and Gold (1962a, b).

the evolution of the system is given by

$$E = E_0 = \text{const}$$
$$\theta = \omega t - \delta \tag{2.3}$$

where δ is some constant angle.

The torus has coordinates x and y which vary between 0 and 1. This should be considered the phase space of the system since, in the dynamics we shall give, the coordinates (x, y) are sufficient data to determine the motion for all time. The time evolution occurs in discrete steps, and each step is taken to be

$$x' = x + y, \bmod 1$$
$$y' = x + 2y, \bmod 1 \tag{2.4}$$

(both radii of the torus are 1). This flow ("automorphism," strictly speaking) is mixing. That is, if μ is Lebesgue measure of the torus we consider two measurable subsets A and B of the torus. Let the transformation described by equation (2.4) be designated φ and its nth iterate φ^n. Then

$$\lim_{n\to\infty} \mu(A \cap \varphi^n B) = \mu(A)\mu(B) \tag{2.5}$$

That this is true for all measurable sets A and B is the mixing property. The harmonic oscillator dynamics is not mixing since every subset of phase space is simply rotated about with no change in shape. For many properties of the flow of equation (2.4), as well as proofs that it is mixing, see Arnold and Avez (1968).

Entropy for either of these two systems is defined as follows. The phase space is devided into cells or grains Δ_i, $i = 1, \ldots,$ of equal measure. The entropy $S(A)$ of a set A is then given by

$$S(A) = -\sum_i p_i \log p_i, \qquad p_i = \mu(A \cap \Delta_i)/\mu(A) \tag{2.6}$$

Increase of S in the course of time (as A evolves) is indicative of progress towards equilibrium. In this respect the flow on the torus behaves well: For reasonable Δ_i, etc., equilibrium is rapidly attained. For harmonic oscillators it is never attained and somewhat artificial means must be employed to force a kind of equilibration. We consider not a single oscillator, but a collection of them with varying frequencies distributed according to $P(\omega)$. When examining the entropy, however, we shall not look in the many-dimensional phase space, but rather the projection of all these oscillators on a single E-θ plane. With this device, if we start the collection according to some distribution in θ and E as well as ω, the averages of θ and θ^2 (denoted by angle brackets) behave as follows:

$$\langle \theta(t) \rangle = \langle \omega \rangle t - \langle \delta \rangle$$
$$\langle \theta^2(t) \rangle = \langle \omega^2 \rangle t^2 + \langle \delta^2 \rangle - 2\langle \omega \rangle \langle \delta \rangle t \tag{2.7}$$

where it is assumed that δ and ω vary independently. A reasonable criterion for equilibrium is

$$\langle [\Delta\theta(t)]^2\rangle \equiv \langle [\theta(t) - \langle\theta(t)\rangle]^2\rangle = 2\pi \tag{2.8}$$

and by equation (2.7) this will take place for τ such that

$$\tau = (2\pi/\langle(\Delta\omega)^2\rangle)^{1/2} \tag{2.9}$$

if it is assumed that $\langle\delta\rangle$ and $\langle(\delta)^2\rangle$ are small (otherwise τ is smaller).

The "experiment" carried out on these models is the following: Two regions of phase space are defined; we require the initial state of the system to be in one of these and the final state in the other. Thus we consider the evolution of a region of phase space and only keep that part of the original region which ends up in another prescribed region. [This prescription replaces the boundary-value problem of Schulman (1973), which is extremely difficult to work with for anything beyond harmonic oscillators.] At some time during the evolution of the system it is subject to a time-reversal-violating perturbation. For the oscillators this takes the form of an instantaneous, totally effective frictional force. If the perturbation takes place at time T_1, its consequences can be summarized by

$$\begin{aligned} q(T_1^+) &= q(T_1^-) \\ p(T_1^+) &= 0 \end{aligned} \tag{2.10}$$

where $f(T_1^\pm)$ means the limit of f as t approaches T_1 from above (+) or below (−). For flow on the torus the perturbation takes the form

$$\begin{aligned} x' &= x \\ y' &= \alpha y, \qquad 0 < \alpha < 1 \end{aligned} \tag{2.11}$$

To discuss the sense in which the operation (2.11) is time-reversal violating it is convenient to use matrix notation. Position on the torus is represented by a column vector ξ with x in row 1, y in row 2. The transformation (2.4) is the matrix M and (2.11), R. These are

$$M = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}, \qquad R = \begin{pmatrix} 1 & 0 \\ 0 & \alpha \end{pmatrix} \tag{2.12}$$

where it is understood that the position vector is always taken modulo 1.

Since the torus, as a phase space, does not naturally yield itself to a description in terms of position and momentum, it is not immediately clear what the time-reversal operation ought to be. A time-reversal operator, T, is a map of phase space into itself. Since $T^2 = 1$, it preserves area. In order that M be T invariant we require (Schulman, 1972)

$$T = MTM \tag{2.13}$$

If all matrix elements are integers, the modulo 1 operation can be taken at the end so that the multiplication in equation (2.13) can be taken to be ordinary matrix multiplication. The most general form of T satisfying equation (2.13) and having integer matrix elements is

$$T = \begin{pmatrix} a & \frac{1}{2}(a+b) \\ \frac{1}{2}(a-b) & -a \end{pmatrix} \tag{2.14}$$

where a and b are integers satisfying

$$5a^2 - b^2 = 4 \tag{2.15}$$

Thus we can take[4]

$$T = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \tag{2.16}$$

The operator R does not satisfy $T = RTR$ with the T of equation (2.16)[5]. Hence R violates T invariance and in fact shrinks the occupied region of phase space.

In setting up the boundary-value problem we do not in general demand that the initial and final sets in phase space coincide–indeed for the frictional force on the oscillators this would eliminate most phase points. Rather our symmetry requirement is that both initially and finally the sets in phase space have the same area. This corresponds to the systems' being dispersed by about the same amount both initially and finally so that final and initial entropy is about the same (the imprecision here is due to the variability possible in the definition of the grains $\{\Delta_i\}$).

3. *Numerical Results*

The previous section was concerned with a set in phase space and its time evolution. In our numerical work we took various sets and considered the evolution of randomly selected points uniformly distributed in this set. These began at $t = 0$ and at $t = T_1$ were subjected to the asymmetric perturbations described above. Then they propagated, according to their usual law, for an additional time T_2, when they were required to fall in another fixed set of phase space. That is, all points from the original set that did not fall in the right place at the end were rejected. Randomly selected frequencies for the oscillators were also used. In this way we enforced symmetric boundary

[4] T of equation (2.16) corresponds to $a = 1, b = 1$. Other solutions are $a = 2, b = 4$ and any pair given by the recursion $a_{n+1} = 3a_n - a_{n-1}, b_{n+1} = 3b_n - b_{n-1}$. We have no reason to prefer any particular a and b. Equation (2.16) is an example of Pell's equation. See H. Hasse (1950).

[5] In fact, R will not satisfy $T = RTR$ for any choice of T satisfying $T^2 = 1$.

Figure 1. Entropy as a function of time. 390 oscillators with random (uniformly distributed) frequencies in the range $10 \leqslant \omega < 10.5$ are located in the region of phase space $13 \leqslant E < 16, 0 \leqslant \theta < 1.2$ (radians) at time $t = 0$. At $t = 67$ they are subjected to friction with the consequences described in equation (2.10) of the text. From $t = 67$ until $t = 118$ they again propagate freely. At $t = 118$ they all lie in the region of phase space $0.1 \leqslant E \leqslant 16.1, 0 \leqslant \theta \leqslant 6.2$.

conditions. Generally we took T_1 approximately but not exactly equal to T_2 because exact equality introduced spurious correlations for the oscillators. In our experience it made no difference whether $T_1 > T_2$ or $T_2 > T_1$. The resulting collection of acceptable points was again propagated from $t = 0$ to $t = T_1 + T_2$ and its entropy calculated [equation (2.6)] as a function of time. So long as the number of acceptable points was several times the number of grains, the behavior of the entropy was not sensitive to the grain size.

In Figure 1 is shown the entropy as a function of time for oscillators starting in a small region of phase space but allowed to end almost anywhere (exact parameters in the figure caption). The evolution is therefore essentially that of an initial-value problem. The entropy behaves as expected. Within a few "seconds" equilibrium is reached and the entropy fluctuates about a value determined by the number of grains and the number of acceptable points (if G is the number of grains and N is the number of acceptable points and $N/G \gg 1$, then $\langle S \rangle = \log G - G/2N$): Prior to the frictional force (at $t = 67$) there is no indication (in S) of the impending perturbation. Subsequently, of course, the entropy increases and the system returns to equilibrium (we cannot account for the mysterious fluctuation in returning to equilibrium after $t = 67$).

When the size of the terminal region in phase space is reduced an entirely different situation emerges. In Figure 2 the oscillators have an initial region of phase space similar to that used for Figure 1. But now they must find their way into a fairly small region of phase space at the end. They never even get close to equilibrium (the entropy staying below $\log G - G/2N$). Moreover, immediately before the friction at $t = 67$, the entropy drops precipitously in anticipation. This is a confirmation of the calculations of Schulman (1973)

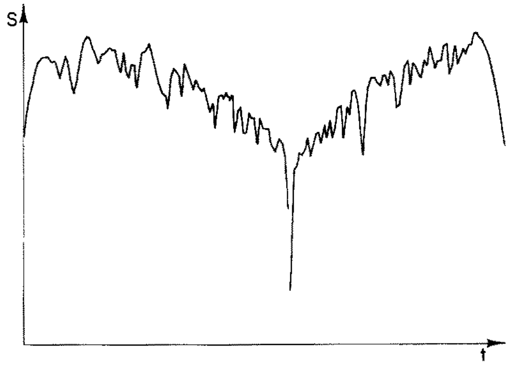

Figure 2. Entropy as a function of time. 258 oscillators with the same frequency distribution as those in Figure 1, begin in the region $15 \leqslant E < 15.25$, $0 \leqslant \theta < 1.2$ at time $t = 0$. At $t = 95$ friction acts. At $t = 172$ they are required to be in the region $5 \leqslant E < 5.75$, $0 \leqslant \theta < 1$.

in which the oscillators were found not to develop an arrow. (In the next section we shall give some general arguments accounting for this anticipatory effect.)

In Figure 3 is shown the entropy as a function of time for the flow on a torus, starting at $t = 0$, with a perturbing projection at $t = 59$. Finally at $t = 98$ the system is required to be in a region of the torus of the same size as that in which it started.

This system has an arrow. The entropy drops, as it should after the perturbation and just before the end. But, despite the end point boundary condition, there is no anticipation at $t = 58$, 57, etc., of the oncoming projection. This system can be said to behave causally because although boundary values have been given, the system acts as if it had only initial conditions.

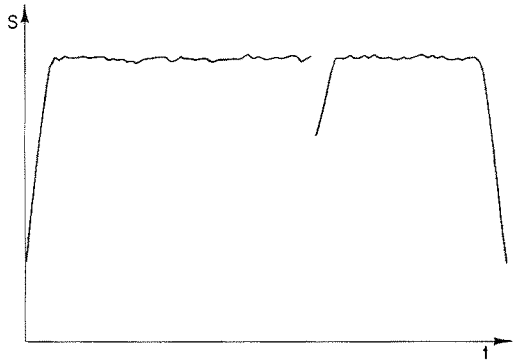

Figure 3. Entropy as a function of time. 204 particles flowing on a torus. They begin ($t = 0$) and end ($t = 98$) in the region $0.35 \leqslant x < 0.45$, $0 \leqslant y < 0.1$. At $t = 59$ they are subjected to the transformation $y' = y/10$, $x' = x$.

4. *General Considerations*

We have considered two model systems, one of which shows an arrow, one of which does not. Which is a better model of the physical world?

Flow on the torus has the advantage of being a mixing system. To us, at least, it seems that this is more characteristic of real systems than the harmoniticity of the systems of oscillators.

However, in another important respect the oscillators are the better model —they have conserved quantity, the energy. We shall now explain the dip in entropy immediately preceding the friction (Figure 2) in terms of the conserved quantity alone. Suppose the oscillators begin in a region (Figure 4a)

$$\begin{aligned} E_i \leqslant E \leqslant E_i + \Delta E_i \\ \theta_i \leqslant \theta \leqslant \theta_i + \Delta\theta_i \end{aligned} \tag{4.1}$$

at $t = 0$. They have some distribution of angular velocities ω. We first follow all points in this region, including those ultimately to be eliminated by the final conditions. If $T_1 \gg \tau$ [equation (2.9)] the system is in equilibrium by time T_1. That is, the points will be uniformly distributed in the annulus

$$E_i \leqslant E < E_i + \Delta E_i \tag{4.2}$$

(see Figure 4b). After the frictional force is applied, they are projected down from the annulus and have $\theta = 0$ or $\theta = \pi$ (Figure 4b). This line of points begins to circle and spread, filling the entire disk (Figure 4d):

$$E < E_i + \Delta E_i \tag{4.3}$$

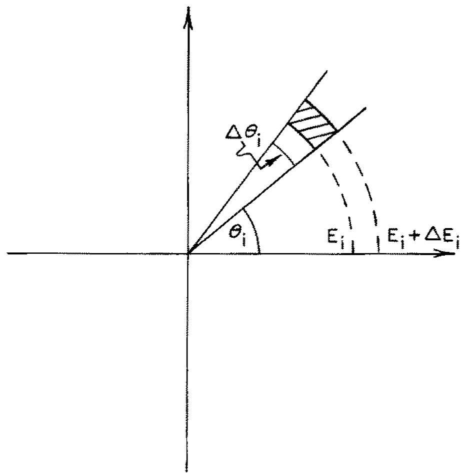

Figure 4a.

Figure 4. (above and next two pages) Progress of oscillators in phase space. See text for explanation. In Figure 4c and 4f, $p = 0$ for all oscillators and the heavy line indicates only which regions in q are occupied.

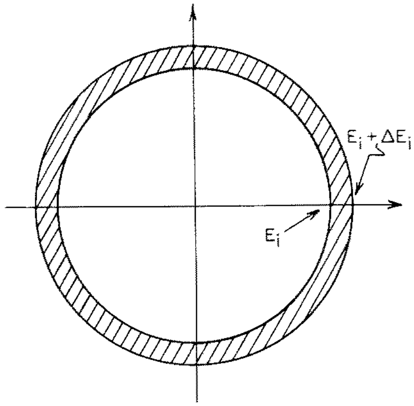

Figure 4b.

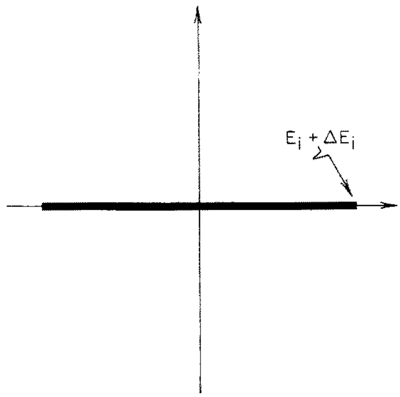

Figure 4c.

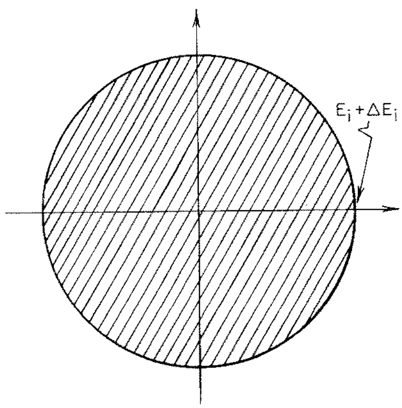

Figure 4d.

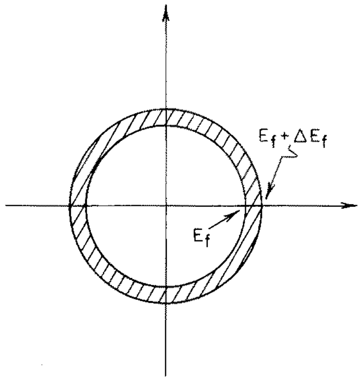

Figure 4e.

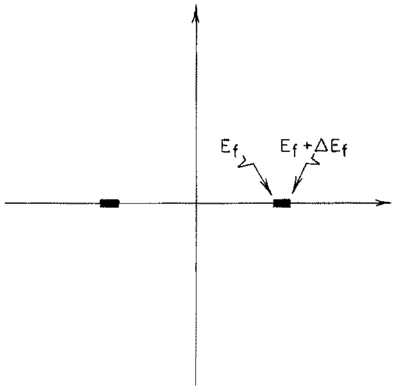

Figure 4f.

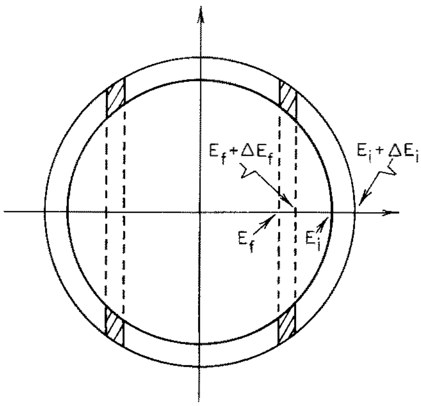

Figure 4g.

But at $t = T_1 + T_2$ we enforce the condition

$$\begin{aligned} E_f \leqslant E < E_f + \Delta E_f \\ \theta_f \leqslant \theta < \theta_f + \Delta\theta_f \end{aligned} \tag{4.4}$$

thereby eliminating many of the points we have been looking at until now. In fact, this means that the points we have kept did not fill the entire disk or inequality (4.3) for $T_1 < t < T_1 + T_2$ but rather, because of conservation of energy, must have been somewhere in the annulus

$$E_f \leqslant E < E_f + \Delta E_f \tag{4.5}$$

(Figure 4e). In particular, this means that for $t = T_1^+$, immediately after the projection, the points were not on the segment pictured in Figure 4c, but rather were on (at most) the intersection of that segment with the annulus of inequality (4.5) (see Figure 4f). But then it follows that immediately before the friction ($t = T_1^-$) the annulus of inequality (4.2) (Figure 4b) was not uniformly filled, but only contained points which projected onto the intersection of the segment and the annulus defined by the final energies (see Figure 4g). Thus for $t = T_1^-$, the system is not in equilibrium and if we follow its development back from T_1 will take some time on the order of τ to equilibrate.

This argument can easily be generalized to arbitrary Hamiltonian systems, including those with better equilibration properties than those of harmonic oscillators.

Suppose the general system is started in some region of phase space, in particular with some more or less well determined energy—a perfectly possible operation in view of the existence of a Hamiltonian. In some time (equilibration or relaxation time) τ the system will have spread uniformly through the energy shell allowed it, say Ω_{E_i}. Now apply some irreversible force—it need not be perfectly effective friction, but only some force that pushes the system to various other energies. Call this operation P. Now the system spreads again, in time filling all regions consistent with energy conservation. However, as for the oscillators, if the final conditions contain energy inequalities, as well they might, only a shell Ω_{E_f} will be filled, which means the operator P can only bring points to within this shell. This implies that prior to the operation of P not all of Ω_{E_i} was uniformly filled but only points in the inverse image of Ω_{E_f} under P, $P^{-1}(\Omega_{E_f})$. We can now state what is required of P for causality: The set $\Omega_{E_i} \cap P^{-1}(\Omega_{E_f})$ must be uniformly distributed in Ω_{E_i}.

Although we have not found any example of an operator P satisfying this condition, we do not know if the condition is a stringent one. It is clearly dependent on the grains taken for the entropy calculation (this is the meaning given to the term "uniformly" used above). One may speculate that it is only satisfied for a particular range E_f, but that this is the energy range that would arise in the initial-value problem.

We have not drawn any conclusions in this section, but only wished to point out pros and cons of each model.

5. *Summarizing Remarks*

We have shown that a certain mixing flow on a torus displays an arrow of time as envisaged in a thought experiment of Gold. The failure of harmonic oscillators to exhibit such an arrow has been reconfirmed. Finally we pointed out that as a model of the world, the flow on the torus had the advantage of being mixing while the poor behavior of the oscillators (as far as getting an arrow is concerned) may arise not from their notorious harmoniticity, but rather may be characteristic of all Hamiltonian systems.

References

Arnold, V. I. and Avez, A. (1968). *Ergodic Properties of Classical Mechanics*, Benjamin, New York.

Gold, T. (1962a). *Recent Developments in General Relativity* Pergamon, London; see also (1962b) *American Journal of Physics*, **30**, 403.

Harrison, E. R. (1974). *Physics Today*, **Feb**, 30.

Hasse, H. (1950). *Vorlesungen über Zahlentheorie*, Springer-Verlag, Berlin.

Schulman, L. S. (1972). *Annals of Physics*, **72**, 489.

Schulman, L. S. (1973). *Physical Review D* **7**, 2868.