Journal of Statistical Physics, Vol. 9, No. 4, 1973

Entropy Changes for Steady-State Fluctuations

R. Landauer[1]

Received June 25, 1973

The dissipative steady state far from equilibrium and subject to a slow modulation of external parameters is analyzed. It is shown that the time-integrated energy dissipation consists of three terms. The first of these is irreversible and consists of the time-integrated dissipation of the sequence of exact steady states defined by the externally controlled parameters traversed during the modulation. The second term is reversible and reflects the fact that the dissipation of the time-dependent modulated system, as calculated in a macroscopic way from ensemble averages, is not the same as the dissipation of a sequence of exact steady states. The third term is also reversible and relates to the ensemble dispersion in changes in stored energy during the modulation. If the system has a single degree of freedom and narrow fluctuations, then these fluctuations can be characterized by an effective temperature T_N. The third term can then be shown to be equal to $T_N\, dS$, where S is the entropy calculated from the distribution function by the usual definition.

KEY WORDS: Entropy; steady state; fluctuations; dissipative systems; noise temperature.

1. INTRODUCTION

This paper will present a generalization of the second law of thermodynamics, $dS = dQ/T$, to some systems which are very simple in some ways, but can be *very far* from equilibrium. Our introductory material relates the subsequent details to earlier work in each of two fields, and tries to make our motivation evident. First of all we will relate our results to work in recent years on the steady state in dissipative systems far from equilibrium. Second, we will

[1] IBM Thomas J. Watson Research Center, Yorktown Heights, New York.

relate the work to a more narrow community of papers, stemming largely from this author's laboratory, and dealing with the physical limits of the computational process, i.e., the *interaction of* information streams.

The relationship presented here was first put forth in detail in an earlier paper.[1] In the present paper we would like to make that material clearer, and correct an error in the earlier discussion. While the earlier paper is believed to be entirely correct in its analytical details, there was a point of confusion in the accompanying verbal interpretation. Furthermore, we shall not try to cover the full range of cases discussed in the original paper.

1.1. Dissipative Systems Far from Equilibrium

There has been a great deal of concern in recent years with the behavior of dissipative systems which are far from equilibrium and exhibit some kind of organized behavior. A conference[2] was devoted to these topics and showed a diversity of approaches and motivations leading to somewhat related descriptions. Much of the work in this field has been concerned with the construction of analogies between distribution functions characterizing nonequilibrium systems and the more familiar $\exp(-\beta H)$ characterizing equilibrium systems. The cited conference proceedings and a review paper currently in preparation,[3] as well as a forthcoming book,[4] provide systematic expositions of these concepts, with an emphasis on cooperative phenomena. The work in Refs. 2–4 is distinguished from much of the earlier work in irreversible statistical mechanics, and also from the early work of Prigogine and collaborators, by its emphasis on the details of the distribution function, which in turn depend on the fluctuations in the system. Later work by Prigogine and collaborators has, however, been concerned with fluctuations and distribution functions.[5,6] The cited discussions[2–4] concentrate on systems which either exhibit detailed balancing, even in the absence of equilibrium, or which can *easily* be manipulated into a form where this is true. It is still unclear (to this author, at least) whether this is a terribly serious restriction, permitting easy characterization of only a few particularly fortunate cases, or whether the techniques are in fact a widely applicable tool.

It is important to note that the use of the macroscopic equations for the time development of a system, without regard for fluctuations, does not permit us to discriminate between metastability and absolute stability; a system without fluctuations will just stay in a metastable state once it is placed there. As an example, consider the tunnel diode circuit of Fig. 1. This circuit contains a battery, a resistance, and a tunnel diode in series. For the purposes of this discussion the details of tunnel diode physics are not relevant. A tunnel diode has an $i(V)$ characteristic, exhibiting a negative resistance ($di/dV < 0$) as shown in the solid curve of Fig. 1(b). The dashed line in

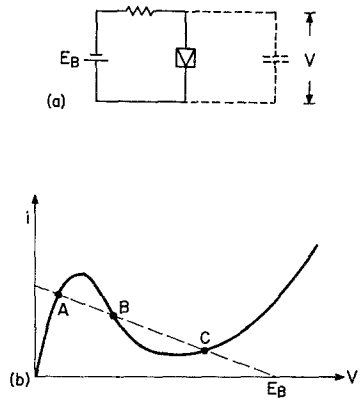

Fig. 1. Tunnel diode fed through series resistance R. V is the potential across the diode. A and C are locally stable, B is unstable.

Fig. 1(b), the "load line," gives the current through the resistor, for a fixed battery voltage E_B, as a function of the voltage V at the junction of resistor and tunnel diode. The two characteristics intersect in three points; two of these are points of relative stability, and one is unstable. These conclusions about stability result from the time-dependent macroscopic equations of motion of this system, which in turn depend on the energy-storing degrees of freedom. In the example shown there is only one degree of freedom, the capacitance of the tunnel diode.

Once the macroscopic kinetics of a system is understood it is not hard to answer questions about *local* stability or instability. To do that, one only has to examine the *linearized* equations of motion for small deviations from the state under consideration and must ask whether such deviations grow or decay. It is not clear to this author why there is such a strong concern elsewhere with stability criteria related to entropy or entropy generation. These criteria in the final analysis depend on the same macroscopic equations of motion.

To decide, however, which of the two states of relative stability, A and C in Fig. 1, is really favored, we have to go beyond the macroscopic equations of motion. Once we have a description of the stochastic behavior of the system we can find the steady-state distribution function, giving the relative probability of finding the system near A versus the probability of finding it near C. We can, however, also go beyond that and describe the relaxation rate toward that steady-state distribution. This depends upon the rate of transition between A and C over the "barrier" at B, much as in a thermally activated atomic jump over a saddle point between favored lattice sites. These

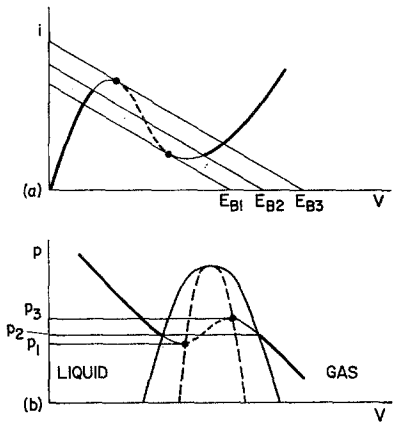

Fig. 2. (a) Tunnel diode characteristics, showing favored steady states by a heavy line. The lighter solid line shows states which are only metastable, while the dashed part of the i–V characteristic shows states which are unstable. (b) The very similar situation for a liquid–gas isotherm.

general points have been understood in connection with electronic systems[7,8] for over a decade.

One can thus arrive at a picture of stability, metastability, and instability for the tunnel diode circuit which is illustrated schematically in Fig. 2. Figure 2 shows behavior under varying battery voltage E_B and is only intended as a qualitative illustration, not as a result of exact calculations. Figure 2 stresses the similarity to a gas–liquid isotherm.

The circuit of Fig. 1, exactly as sketched, does not strictly satisfy detailed balance, as pointed out in the appendix. A circuit which is readily tractable by the methods we have discussed[2–4] is obtained by replacing the resistance by a device which, just like the tunnel diode, passes one whole electron at a time. This could then, for example, be another tunnel diode, or a thermionic diode, or an insulating layer which transmits by tunneling. Since these latter circuits, however, lead to circuit diagrams which are less familiar and less suggestive, we have invoked the resistor in the abbreviated discussion given above. More details concerning this will be found in the appendix.

The relationship $dS = dQ/T$, to be proven for steady-state dissipative systems under a slow change of parameters, will extend the analogies[2–4] and give them a more physical and thermodynamic significance. The relationship $dS = dQ/T$ tell us, for example, that entropy reduction in equilibrium is accompanied by a heat transfer to the thermal environment. We shall show that for a very limited variety of dissipative systems there is a similar relation.

If we vary the parameters of the system slowly and in such a fashion as to narrow the system's range of allowed behavior, then there will be an associated reversible heat flow to the environment.

1.2. INFORMATION PROCESSING

Systems with two states of local stability are, of course, the components out of which computers are built. A concern with the ultimate physical limitations of computers has been one of the motivations for the studies discussed in this paper. Information handled in computers, in biological systems, or even by pencil and paper inevitably utilizes real physical degrees of freedom. Thus information is not a disembodied philosophical or mathematical entity, and is subject to physical restrictions. The attempt to understand these restrictions is still in its infancy.[9,10] A chief concern has been to find the minimum energy expenditure required in the computational process. This is most easily discussed in systems in which information can be held in locally stable states without requiring continuous energy dissipation and where energy dissipation is only required when information is changed. Systems of this sort are exemplified by magnetic cores which can be magnetized in one of two directions, or by a particle in a bistable potential well.

In the existing discussions[9,10] it has been shown, through phase-space arguments, that operations which throw away information, i.e., operations which do not allow a deduction of the input from the output, require energy dissipation. The amount of the dissipation depends upon the particular logical function being executed, but for typical elementary logic functions is of order kT. This author has furthermore argued that operations which throw away information are essential to a computer. In a very remarkable recent paper Bennett[11] has shown that, in fact, computers need not throw away any information. Let us here briefly try to describe the spirit of his argument. Consider a two-input, one-output function, such as the logical "and", whose output is "1" if and only if both inputs are "1." Clearly it is logically irreversible, and therefore[9,10] also physically irreversible and dissipative. Now, however, let us supplement the needed "and" output by two other outputs which simply replicate the inputs. Or, to put it more generally: provide enough added output to permit us to go back in a one-to-one fashion from output to input. Such a one-to-one mapping of input into output can be performed by a dissipationless mechanism, simply following along its dynamic trajectories.

Now, however, we are left with extra, unnecessary, outputs. These cannot be thrown away; that requires energy dissipation. The extra outputs must therefore be fed into a shift register[2] or some other sort of memory. At

[2] A shift register is a storage device which, as its name implies, shifts information along a linear track as new information is fed in from one end.

the end of the computation we will, of course, have a good deal of unnecessary intermediate results filling shift registers. In earlier discussions[12] it was assumed that these shift registers would eventually have to be erased, to clear the computer for its next use, and that therefore this whole scheme served only to postpone the dissipation in time, rather than to avoid it. Unfortunately, this was an erroneous conclusion, resulting from the fact that we were considering "driven computers," as Bennett calls them. These are computers which always move forward from logical antecedents to logical consequences at a (roughly) predictable rate. Indeed, all real electronic computers are of this sort. But computers need not be "driven," they can be allowed to be physically reversible. Once we allow real physical reversibility then the unnecessary intermediate results saved in shift registers need not be erased. After arriving at the end of the computation we copy the desired results. We then let the computer run backward, and in this process the shift registers unwind and are cleared. Since the whole computation consists of steps each of which is physically and logically reversible, the computer will simply run backward to its original input state. Naturally, of course, the forward motion as well as the backward motion requires a little energy or driving force to overcome viscosities. These energies, however, can be very much smaller than kT per logical step.

Actual computing systems, however, typically use dissipative components, requiring energy dissipation not only for the forward motion, but simply to hold results. A long-standing speculation exists[12] that, "...in the steady-state dissipative device the dissipation per switching event is at least as high as in the [conservative devices], and that this dissipation per switching event is supplemented by the steady-state dissipation." In the subsequent discussion we shall see that this may be somewhat too simple a viewpoint.

The generalizations of the second law of thermodynamics which are the concern of the remainder of this paper result from the attempt to become more analytical about the speculation given above, by relating entropy changes in steady-state systems to heat exchanges with the reservoir. Unfortunately, at this time a complete characterization of the dissipative systems comparable to that existing for devices holding information without dissipation does not exist.

2. OUTLINE OF THE THEORY

We are typically concerned with systems with one degree of freedom, such as the capacitance in the tunnel diode circuit of Fig. 1. This degree of freedom is connected to a circuit which in the steady state brings the capacitance to a state of preferred charge. There are also fluctuations in the circuit tending to drive the capacitor charge q away from its preferred value q_0. The

balance between the restoring forces and the fluctuations is typically represented by a Fokker–Planck expression for the ensemble flux j_q along the q axis:

$$j_q = \rho(q)\, v_q - D(\partial\rho/\partial q) \tag{1}$$

ρ is the probability distribution for ensemble members, v_q is the restoration velocity toward the nearby stable or metastable steady state, and D describes the diffusive motion of ensemble members. Thus D describes the tendency of ensemble members originally at the same value of q to separate from each other. The Fokker–Planck equation requires justification, which we will not provide, though it will turn out that our subsequent argument does not absolutely need to go back to the Fokker–Planck equation. In particular, Eq. (1) assumes, of course, that ensemble members make transitions only to nearby values of q. In the case of the tunnel diode the Fokker–Planck equation is validated in the original paper[8] on that device, and that discussion is supplemented in the appendix. Problems with the Fokker–Planck equation have been discussed in detail by van Kampen.[13] In fact, our tunnel diode circuits obey a master equation which *cannot* validly be approximated by a Fokker–Planck equation, except in the vicinity of a steady state. That, however, is where we shall really need the solution.

There is another subtle question which concerns the detailed form of the Fokker–Planck equation. Is it as written in Eq. (1), or should the diffusion coefficient appear under the differentiation sign? Both forms of the diffusion current can actually be correct, and are used in the literature, but they correspond to, and require, different definitions for v_q. This has been discussed in detail in the appendix of Ref. 9 and also by Manning.[14]

In one dimension Eq. (1) is trivially integrated to yield the steady-state distribution

$$\rho = A \exp\left(\int v_q\, dq/D\right) \tag{2}$$

with a normalization constant A. The thermal equilibrium distribution function $\rho \sim e^{-U/kT}$ can be regarded as a special case of Eq. (2), obtained by utilizing the Einstein relation in the form $v_q = -(D/kT)(dU/dq)$.

The velocity of restoration v_q toward a preferred state q_0 will go to zero at that state. The integral in Eq. (2) will therefore be a maximum at that point. It will fall off in either direction, as the restoration velocity $v_q \simeq -\alpha(q - q_0)$ increases in magnitude as we move away from q_0. Thus if α and D are very slowly varying functions of q over the range in which ρ is appreciable, then Eq. (2) becomes

$$\rho = A \exp[-\alpha(q - q_0)^2/2D] \tag{3}$$

The fact that α and D are actually not constants will give us a more asymmetric distribution function than shown in Eq. (3), and the problems caused

by such asymmetries are the main concern of Section 5 and 6. The energy $U(q)$ stored in the system can be approximated over this same narrow range by

$$U(q) = U(q_0) + (q - q_0)\, U' + \tfrac{1}{2}(q - q_0)^2\, U'' \tag{4}$$

If we once again neglect higher-order asymmetry, we can rewrite Eq. (3) in the form

$$\rho = A \exp\left\{-\frac{\alpha}{2D}\frac{2}{U''}\left[U(q) - U(q_0) - (q - q_0)\, U'\right]\right\} \tag{5}$$

Equation (5) is the thermal equilibrium distribution function for a system at a temperature $kT_N = DU''/a$ and with a fixed force $U'(q_0)$ applied to the system to displace it to q_0. Thus the distribution function for the dissipative system is characterized by an effective temperature. This temperature is determined by the noise sources and is proportional to D, and is not in general the ambient temperature unless we are dealing with noise sources obeying the Nyquist relationship. Our effective temperature is essentially the "noise temperature" of the electrical engineering literature.

Since the steady-state distribution simulates an equilibrium distribution, we can assign it the same entropy

$$S = -k \int \rho \log \rho \, dq \tag{6}$$

assigned to the thermal equilibrium distribution. For the thermal equilibrium distribution, as we slowly change parameters (e.g., q_0 or T) we have $dS = dQ/T$. Now two of these quantities, dS and T, also seem to characterize a shifting steady-state distribution, as we have just seen. This leaves us with the job of finding a proper interpretation for dQ in the steady-state case.

For subsequent use let us recast the relationship $dS = dQ/T$ *for the equilibrium case* in an alternative form. From Eq. (6) we find

$$\delta S = -k \int (\delta\rho) \log \rho \, dq \tag{7}$$

where we have now written δ to indicate the small changes arising from the shifting state, to prevent confusion with the integration over dq. The term $\log \rho$ in the integrand on the rhs of Eq. (7) can be replaced by using $\rho = Z^{-1}e^{-\Delta/kT}$, where $Z = \int e^{-\Delta/kT}\, dq$, and

$$\Delta = U - U(q_0) - (dU/dq)_{q_0}(q - q_0) \tag{8}$$

is the relative energy in the presence of the bias force, illustrated in Fig. 3. Thus

$$\log \rho = \log Z^{-1} - (\Delta/kT) \tag{9}$$

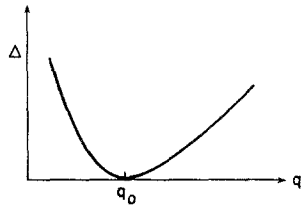

Fig. 3. The relative energy, in the presence of the bias force, about the preferred state. The figure intentionally emphasizes the possible asymmetry.

When we substitute this expression into Eq. (7) $\log Z^{-1}$ drops out in the integration. Thus

$$\delta S = (1/T) \int \delta\rho \, \Delta(q) \, dq = (1/T)[\delta\langle U\rangle - (dU/dq)_{q_0} \, \delta q_0] \tag{10}$$

Here $\langle U\rangle$ is the ensemble average of the energy. The term $(dU/dq)_{q_0} \, \delta q_0$ is just the energy change of the preferred state, i.e., the state at which dU/dq equals the applied force. We shall use U_F to denote the excess of energy $\langle U\rangle$ over the energy at the preferred point $U(q_0)$, emphasizing that it is an energy associated with the fluctuations of the distribution. Then Eq. (10) can be written

$$T \, \delta S = \delta U_F \tag{11}$$

Since Eq. (11) has been derived from the form of the distribution function without explicitly invoking the second law of thermodynamics, it should hold equally for the nonequilibrium case as long as the two distribution functions are really as close to each other as we have suggested. (The fact that they are really not that close is the reason for much of the subsequent discussion.) At this point, however, Eq. (11) is only a relationship between quantities characterizing a shifting distribution; it does not tell us anything about heat exchanges with the reservoir.

In the equilibrium case that connection is made very easily through the first law of thermodynamics. As the point q_0 is shifted the internal energy changes by $\delta\langle U\rangle$. The externally applied force supplies an energy $(dU/dq)_{q_0}\delta q_0$. The difference between these two is δU_F (this difference is analogous to the more customary expression $dU - p \, dV$), and this must come from the only other possible energy source, the thermal reservoir. Hence $\delta U_F = \delta Q$, and together with Eq. (11) this demonstrates $T \, \delta S = \delta Q$ for the equilibrium case. In the nonequilibrium case, however, we cannot simply invoke conservation of energy, in view of the background of steady-state dissipation. In the next section we shall attempt to remedy this shortcoming and supply the necessary relationship.

3. GENERALIZATION OF THE FIRST LAW OF THERMODYNAMICS

The title of this section must be taken with a grain of salt. The first law does, of course, apply directly as it stands to dissipative systems. We are here instead referring to a generalization of $\delta U_F = \delta Q$ for the specific purposes of the last section.

Conservation of energy, in its normal sense, does apply to the dissipative system under consideration. Thus there is energy supplied externally, e.g., by batteries in the case of an electrical circuit. This must equal the sum of the energy dissipated in the same period and the change in stored energy. Thus

$$\int (\text{power supplied})\, dt = \int (\text{heat dissipation})\, dt + \delta U \tag{12}$$

Since this equation applies to every system, it applies also to ensemble averages.

Now consider the macroscopic system without fluctuations, i.e., the system as characterized by ensemble averages for voltages and currents. The macroscopic circuit equations between these also satisfy the law of conservation of energy. (If this is not considered an obvious fact, then a more detailed argument can be found in the original treatment.[1]) This yields

$$\int (\text{power supplied})_0\, dt = \int (\text{heat dissipation})_0\, dt + \delta U_0 \tag{13}$$

where the subscript zero denotes the fact that these quantities are calculated in the usual macroscopic way from ensemble averages, e.g., heat dissipation in a resistor is now taken as $\langle i \rangle \langle V \rangle$, rather than as $\langle iV \rangle$ in Eq. (12). The term δU_0 in Eq. (13) is $\langle dU/dq \rangle \langle \delta q \rangle$. Now if our steady-state distribution rigorously simulates a Boltzmann distribution, then $\langle dU/dq \rangle = (dU/dq)_{q_0}$, where q_0 is the point at which ρ has a maximum. (We shall return to a discussion of the accuracy of the simulation in the next two sections.) In case it is not obvious to the reader that $\langle dU/dq \rangle = (dU/dq)_{q_0}$ for the Boltzmann distribution, this can be shown through a simple integration of $\int \rho (dU/dq)\, dq$. The identification between ensemble averages and quantities obeying the macroscopic circuit equations only makes sense if the circuit equations have single-valued solutions, or if only one of several possible multivalued solutions is involved. If in Fig. 1 solutions near A and C are both represented in the ensemble, then the ensemble averages correspond to averages over suitably weighted solutions of the circuit equations. If we are dealing with such a true steady state, where the various branches are in equilibrium with each other, our equations are still applicable, but their significance is not yet entirely clear.

Now take the difference between Eqs. (12) and (13). The power supply terms will cancel exactly, since the voltage supply (or current supply, if preferred) is not a fluctuating quantity. The result is then

$$\langle\text{energy given up by reservoir}\rangle - (\text{energy given up by reservoir})_0 = \delta U_F \tag{14}$$

Here δU_F is $\langle \delta U \rangle - \langle dU/dq \rangle \langle \delta q \rangle$. The replacement of $\langle dU/dq \rangle$ by $(dU/dq)_{q_0}$ is justified only if the distribution is rigorously a Boltzmann distribution. Hereafter we will replace the phrase "energy given up by reservoir" by δQ. The energies given up by the reservoir invoked in Eq. (14) have the same magnitude, but differ in sign, from the heat dissipation terms invoked in Eqs. (12) and (13). Note that for an unmodulated steady state $\delta U_F = 0$, and therefore $\langle \delta Q \rangle = (\delta Q)_0$. This means that for such a genuine steady state, not subject to any shift, the heat dissipation calculated in a macroscopic way by multiplying ensemble averages is in fact the exact heat dissipation, and no extra term has to be added to allow for fluctuations.

We must now emphasize a point which was unfortunately missed in the original treatment.[1] Equation (14) when applied to the states of a system subject to a change with time refers to the behavior of the distribution function *for that time-dependent system.* If the time-dependent modulation of system parameters is slow, then the time-dependent distribution functions will be close to those of the steady-state system, but they will not be identical. Consider for example the circuit of Fig. 1 in the presence of a slowly rising battery voltage E_B. This means that the charge q on the capacitor is also rising slowly. This charging current dq/dt must be supplied by a difference between the resistive current and the diode current. Thus these currents are perturbed away from their steady-state values, and the associated energy dissipation in each component is also perturbed by amounts proportional to the changes in current, and therefore proportional to dE_B/dt. (If our modulated circuit remains close to thermal equilibrium, or alternatively if all the resistances are linear, then the change in total circuit dissipation can be shown to be second order in dE_B/dt. We shall, however, continue to keep the more general case in mind.) Thus the change in energy dissipation is small if the rate of modulation is slow, but the time over which these perturbations exist increases inversely with the modulation rate. *Therefore the integrated change in power dissipation cannot be neglected.*

The distribution functions in the presence of the modulation will hereafter be called the "lagging" states, and will be denoted with a subscript l. We use the word "lagging" since these are states in which the actual distribution function lags behind the steady-state distribution function for capacitive charge defined by the currently valid, externally imposed circuit parameters.

Such lagging distribution functions caught in slowly shifting potential wells have been studied in an erlier paper[15] oriented to very different purposes. In contrast to the lagging states, the subscript ss will denote the exact distribution function of the steady-state system. In the presence of the time-dependent modulation ss will denote a sequence of distribution functions obtained by selecting at each instant of time the exact steady-state distribution function for the modulation parameter prevailing at that exact instant.

In the case of the dissipative terms we have stressed that small dissipation differences between the lagging state and the steady state can be integrated over a long time, and therefore cannot be neglected. The terms $\langle \delta U \rangle$ and $\langle dU/dq \rangle \langle \delta q \rangle$ are simpler. As the modulation rate is decreased and the two kinds of distribution function approach each other we can approximate

$$\langle \delta U \rangle_l - \langle dU/dq \rangle_l \langle \delta q \rangle_l = \langle \delta U \rangle_{ss} - \langle dU/dq \rangle_{ss} \langle \delta q \rangle_{ss} = \delta U_{F,ss} \quad (15)$$

From Eq. (11) we know, however, that $\delta U_{F,ss}$ is $T_N \, \delta S$, where T_N is the temperature describing the steady-state distribution function. Thus Eq. (14) becomes [invoking the definition of δQ that follows Eq. (14)]

$$\langle \delta Q \rangle_l - \delta Q_{0,l} = T_N \, \delta S \quad (16)$$

The entropy change is now related to the energy taken from the reservoir, corrected for the energy exchange expected from the macroscopic circuit equations. Reference 1 labeled this difference δQ_F and derived this relationship correctly, but in some of its verbal interpretation erred, since at that time adequate distinction between the "l" states and the "ss" states was not made. We shall elaborate on this important point in the next section.

4. REVERSIBLE AND IRREVERSIBLE HEAT FLOW

For the purposes of this section we will recast our basic theorem, Eq. (16). We have stressed that

$$\delta Q_{0,ss} - \delta Q_{0,l} = \delta W_{l,ss} \quad (17)$$

does not vanish, but approaches a fixed limit, called $\delta W_{l,ss}$, as the modulation rate is decreased. With this notation we can rewrite Eq. (16) as

$$\langle \delta Q \rangle_l = \delta W_{l,ss} + T_N \, \delta S + \delta Q_{0,ss} \quad (18)$$

Thus the actual heat flow in the presence of the modulation consists, first of all, of the last rhs term, i.e., the dissipation predicted from the sequence of the steady states defined by the time dependence of the modulation. This is an irreversible term. In other words, if the modulation takes our system from

state A to state B and back again, this term has the same sign and magnitude. By contrast, the other two terms are reversible. Part of the reversible heat exchange comes from $\delta W_{i,\mathrm{ss}}$, a simple macroscopic correction allowing for the fact that the modulated system has a different dissipation. The sign of this term, however, will depend on the direction in which the system is changing. Finally, there is the $T_N\,\delta S$ term, telling us that there is an additional reversible heat change which, as in equilibrium, depends on the increase or decrease of the spread of the distribution function.

5. THE DEVIATION FROM THE EXACT BOLTZMANN DISTRIBUTION

In Section 3 we assumed that the distribution function for the steady state was *identical* to that obtainable in a thermal equilibrium situation. There is clearly an approximation here. Typically, for example, v_q and D as invoked in Eq. (2) will both depend on q. Their ratio, in the range of fluctuations, need not correspond to the exact value,

$$\frac{v_q}{D} = -\frac{\mu\,\nabla U}{D} = -\frac{\nabla U}{kT_N} \tag{19}$$

predicted from the shape of the potential, and the effective temperature T_N which characterizes the behavior right at the center of the distribution. In other words, the potential U can have asymmetry, as indicated in Fig. 3, and similarly the distribution given by Eq. (2) can have asymmetry, and these deviations need not be linked to each other. In this section we want to show that as the system under consideration becomes larger and the fluctuations involved become narrower the relative error involved in equating the two distributions becomes negligible.

An equilibrium distribution is of the form $\rho \sim e^{-\Delta/kT}$, where Δ is the relative energy, in the presence of the applied force, as given by Eq. (8) and illustrated in Fig. 3. In the steady-state system ρ can deviate from $e^{-\Delta/kT_N}$, e.g., increase above it for $q > q_0$, and drop below it for $q < q_0$. We have presumably chosen T_N correctly to describe the variation correctly right at q_0, to second order in $(q - q_0)$, and are therefore interested in the effects of a_3, etc., in the following expression:

$$\rho \sim [\exp(-\Delta/kT_N)]\exp[a_3(q - q_0)^3 + a_4(q - q_0)^4 + \cdots] \tag{20}$$

The dominant term will presumably be a_3 and we shall look at its role in some detail. Since Δ varies quadratically, at first, with q about q_0, we can replace $(q - q_0)$ in Eq. (20) by $\Delta^{1/2}$, yielding

$$\rho \sim [\exp(-\Delta/kT_N)]\exp[-\beta(\Delta/kT_N)^{3/2}] \tag{21}$$

to lowest order in β, i.e., to lowest order in the deviations from a strict Boltzmann distribution. Let us consider the behavior of β as the system is increased in size, i.e., as we go to the "thermodynamic limit." Our discussion will invoke some of the specifics of electrical circuits, but hopefully will apply more generally. Admittedly, the real range of applicability of our arguments is uncertain.

We approach the "thermodynamic limit" by imagining the typical sort of circuit we have been considering to be built up by replacing each initial circuit element by a number of identical elements in parallel. Let n denote the number of elements in parallel. Thus we can imagine the value of the load resistance R in Fig. 1 to be divided by n; similarly we have n tunnel diodes in parallel, but all across the unchanged battery voltage E_B .

The dynamics of the system, e.g., the variation of v_q and D in Eq. (2) with q, is essentially determined by the intensive parameter, i.e., by the voltage, which controls the physical details of the tunneling situation. Thus the relaxation rate, $(q - q_0)^{-1}\, dq/dt$, at a given voltage away from the preferred voltage must be independent of n. The diffusion coefficient D describes the buildup of charge fluctuations, $(q - q_0)^2 \sim 2Dt$, that would occur in the absence of a restoring velocity. Each of the n parallel circuit elements contributes independently to these mean-square fluctuations, and thus D is proportional to n. Thus

$$\int \frac{v_q}{D}\, dq = \int_{V_0}^{V} \underset{\downarrow}{\frac{1}{D}}\ \underbrace{\underset{\downarrow}{\frac{1}{q - q_0}}\ \underset{\downarrow}{\frac{dq}{dt}}}\ \underbrace{\underset{\downarrow}{(q - q_0)}\ \underset{\downarrow}{\frac{dq}{dV}}}\ dV \qquad (22)$$

$$(1/n) \quad n^0 \quad n \quad n$$

where the arrows indicate variations with n. For a given voltage deviation $\int (v_q/D)\, dq$ thus varies as n. Furthermore, since $q - q_0$ is approximately proportional to the voltage deviation and the remaining quantities are approximately constant over the range of integration, the integral varies as $(V - V_0)^2$ for a fixed value of n. Thus, letting

$$V - V_0 = \Delta V,\ \rho \sim \exp[-Cn(\Delta V)^2].$$

The range of voltage fluctuations is therefore proportional to $1/\sqrt{n}$. The range of charge fluctuations $q - q_0$ is proportional to both the voltage fluctuations and the size of the capacitor. Thus it is proportional to $\sqrt{n}$. This makes the range of Δ, which is proportional to $\Delta q\ \Delta V$, independent of n, and therefore—not surprisingly—T_N is also independent of n. To understand the role of the modifying multiplier $\exp[-\beta(\Delta/kT_N)^{3/2}]$, consider Eq. (22) and now let one or more of the quantities, such as D, for example, vary as

$$1/D \sim (1/D_0)[1 + \eta(V - V_0)] \qquad (23)$$

where η is independent of n. This then gives us an additional term in the integral which varies as $\eta n \Delta V^3$ and can be identified with $\beta(\Delta/kT)^{3/2}$. Equating these leads to the conclusion that $\beta \sim n^{-1/2}$. In the limit of large n the steady-state distribution therefore approaches the Boltzmann distribution.

While this essentially completes our proof, we shall in the next section examine the fact that the steady-state distribution does differ from the Boltzmann distribution. The surprising result is as follows: If the deviation is taken into account to first order in β, then the relationship $T_N \, \delta S = \delta Q_F$ is still valid.

Equation (21), accompanied by the statement that $\beta \ll 1$, is the basic assumption needed for our theorem, rather than the Fokker–Planck equation. Equation (21) is more general, e.g., it applies directly to a ladder of discrete levels, as discussed in the appendix. It can also be applied to situations in

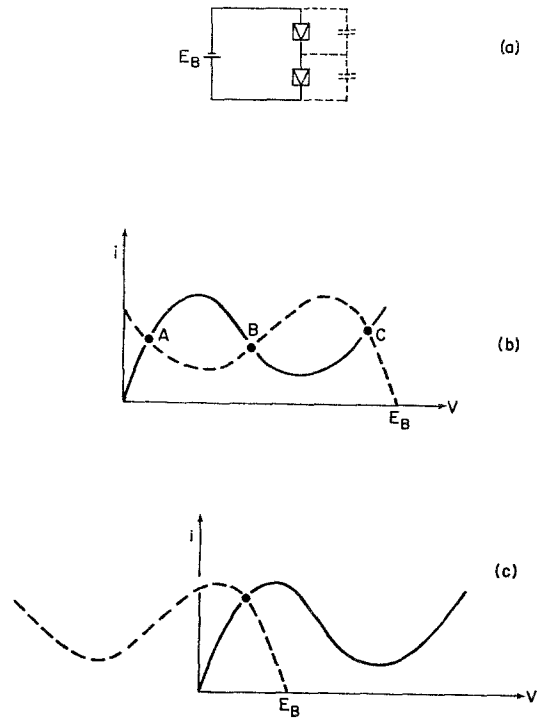

Fig. 4. (a) Bistable tunnel diode circuit, with associated tunnel diode capacitances. (b) Solid line is the current through the lower tunnel diode as a function of the voltage V at the junction between the two diodes. Dashed line gives current through the upper diode as a function of the same midpoint voltage V. (c) Corresponding situation for a lower battery voltage E.

which the random jumps are not small enough to be characterized by a diffusion term.

We have pointed out that if the system is large enough and the fluctuations narrow enough, then the dissipative system and the equilibrium system both become narrow and the distribution functions very similar. The size of the system, i.e., the multiplier n required for this, is a function of the range of fluctuations existing before the circuit elements are placed in parallel. Thus if either the equilibrium system or the dissipative system is near a critical point separating a range of bistability from a range of monostability, then a particular large multiplier is required. Alternatively, a system of *fixed size* brought close to a critical point will reach a stage where the distribution function similarity will break down. (A system with a continuous transition between monostability and bistability is exhibited in Fig. 4.) This is unfortunate since the transition through the critical point is an essential part of some otherwise particularly easily analyzed data processing schemes.[16]

6. EFFECT OF DEVIATIONS

In Eq. (14) we showed that

$$\langle \delta Q \rangle_l - \delta Q_{0,l} = \delta Q_F = \langle \delta U \rangle - \langle dU/dq \rangle \langle \delta q \rangle \tag{24}$$

Using Eq. (8) we can rewrite this, without approximation, as

$$\delta Q_F = \langle \delta U \rangle - (dU/dq)_{q_0} \langle \delta q \rangle - \langle d\Delta/dq \rangle \langle \delta q \rangle \tag{25}$$

We want to compare this to

$$T_N \, \delta S = -kT_N \int \delta\rho \log \rho \tag{26}$$

which can be rewritten, using Eq. (21), as

$$T_N \, \delta S = -kT_N \int \delta\rho [-(\Delta/KT_N) - \beta(\Delta/kT_N)^{3/2}] \, dq \tag{27}$$

Replacing Δ in Eq. (27) through the use of Eq. (8), this becomes

$$T_N \, \delta S = \langle \delta U \rangle - (dU/dq)_{q_0} \langle \delta q \rangle + \beta k T_N \int \delta\rho (\Delta/kT_N)^{3/2} \, dq \tag{28}$$

In thermal equilibrium $\beta = 0$, and the final rhs terms of both Eqs. (25) and (28) vanish, and $T_N \, \delta S = \delta Q_F$. We now want to recognize, however, that in the dissipative steady state as the system becomes large β becomes small, but is nonvanishing. Therefore in comparing the final rhs terms of Eqs. (25) and (28), we will ignore higher-order corrections in β.

Consider first the $\langle d\Delta/dq \rangle$ term, from Eq. (25):

$$\left\langle \frac{d\Delta}{dq} \right\rangle = \frac{\int \frac{d\Delta}{dq} \exp - \frac{\Delta}{kT_N} \exp - \beta \left(\frac{\Delta}{kT_N}\right)^{3/2} dq}{\int \exp - \frac{\Delta}{kT_N} \exp - \beta \left(\frac{\Delta}{kT_N}\right)^{3/2} dq} \tag{29}$$

To first order in β, we can take the normalization integral in the rhs denominator simply as $Z = \int e^{-\Delta/kT}\, dq$. In the rhs numerator of Eq. (29) we take

$$\exp[-\beta(\Delta/kT_N)^{3/2}] = 1 - \beta(\Delta/kT_N)^{3/2} \tag{30}$$

Only the β term from (30) contributes to the integral, in the numerator of Eq. (29), leaving

$$\left\langle \frac{d\Delta}{dq} \right\rangle = -\frac{\beta}{Z} \int \frac{d\Delta}{dq} \left(\frac{\Delta}{kT_N}\right)^{3/2} e^{-\Delta/kT_N}\, dq = -\frac{\beta k T_N}{Z} 2 \int_0^\infty e^{-x} x^{3/2}\, dx \tag{31}$$

The definite integral in Eq. (31) is a gamma function, and therefore the final rhs term of Eq. (25) becomes

$$-\langle d\Delta/dq \rangle \langle \delta q \rangle = (\beta k T_N/Z)\, 2\Gamma(\tfrac{5}{2}) \int \delta\rho\, q\, dq \tag{32}$$

Consider now the other term involved in our comparison, the last rhs term of Eq. (28),

$$\beta k T_N \int \delta\rho (\Delta/kT_N)^{3/2}\, dq \tag{33}$$

To zeroth order in β we have $\rho = Z^{-1} e^{-\Delta/kT_N}$. The change $\delta\rho$ has two sources. First of all the distribution can be shifted in space; additionally its width, i.e., temperature, can change. The latter kind of change is an even function of $\Delta^{1/2}$, and does not yield a net contribution to the integral in (33). A displacement of the distribution in q by δq_0 gives

$$\delta\rho = -q_0 \frac{\partial\rho}{\partial q} = \delta q_0 \frac{1}{kT_N Z} e^{-\Delta/kT_N} \frac{\partial\Delta}{\partial q} \tag{34}$$

If this is entered into (33), we find

$$\beta\, \delta q_0 (1/Z) \int_{-\infty}^{\infty} e^{-\Delta/kT_N} (\Delta/kT_N)^{3/2} (\partial\Delta/\partial q)\, dq = \beta k T_N Z^{-1} 2\Gamma(\tfrac{5}{2})\, \delta q_0 \tag{35}$$

which is equivalent to Eq. (32). Thus to first order in β these terms which represent the effect of deviations from the Boltzmann distribution still obey $T_N\, \delta S = \delta Q_F$.

It is perhaps also relevant to discuss the size of these terms. The main terms in Eq. (25) and Eq. (28), e.g., $\langle \delta U \rangle$, vary as n (n is the number of circuit elements in parallel, as in Section 5) if the shift takes us through a fixed voltage range δV. The terms in (32) and (35) vary as $\beta Z^{-1}\, \delta q_0$ as n is varied. β varies as $(1/n)^{1/2}$, as shown in Section 5. Z varies as the range in q over which the Boltzmann exponential is integrated, i.e., as $n^{1/2}$. δq_0 is the shift in charge, for a fixed voltage shift, and therefore varies as n. Thus $\beta Z^{-1}\, \delta q_0$ varies as n^0 and is therefore small, for large n, compared to the main terms.

7. OVERVIEW

Our basic theorem, Eqs. (16) and (18), has been proven so far for the case of one degree of freedom. A circuit with several degrees of freedom is more complex. There we will have to contend with the fact that normal coordinates can be defined only locally, in the general nonlinear case, through the quadratic parts of the energy variation about a particular point of operation. The local normal coordinates will change their identity as a point of operation is shifted. An appreciably simpler case exists, however, if the total energy can be written as a sum of terms, not necessarily quadratic, one for each of several normal coordinates, which in turn preserve their identity as the point of operation is shifted. This is the case in which we will have normal coordinates q_α and associated forces $\partial U/\partial q_\alpha$ which obey fixed linear relationships to, respectively the charge flow and the voltages at the terminals interfacing the dissipative network. This case includes two important subcases. One is the case where the reactances in the circuit are all linear. The other is where the reactances are all separate two-terminal devices, and are coupled to each other only through the dissipative network.

If we have fixed normal coordinates, we need only assume a distribution of the form (21) for each normal coordinate, with a separate temperature for each normal coordinate. We need not assume that the distribution functions for the various normal coordinates represent uncorrelated probabilities, and need not assume detailed balancing. In that case the discussion in Section 6 applies immediately to each normal coordinate, and $T_N\, dS$ is now replaced in Eqs. (16) and (18) by a sum of terms $T_{Ni}\, dS_i$, one for each normal coordinate.

Even in the cases where Eqs. (16) and (18) or their obvious generalization just discussed do not apply, there is a more general content to our discussion which we want to reemphasize here. The most delicate aspects come from our attempt to show $\delta U_F = T_N\, \delta S$. Equation (18) can, however, be written in the more generally valid form

$$\langle \delta Q \rangle_l = \delta W_{l,\mathrm{ss}} + \delta U_F + (\delta Q)_{0,\mathrm{ss}} \tag{36}$$

Again we can point out that the first two terms on the rhs represent reversible heat flow, while the final term does not. Furthermore, on the right hand side only the term δU_F is sensitive to the details of distribution functions. Equation (36) does not require assumptions about narrow fluctuations or about a single effective temperature. In connection with Eq. (36) it is worth stressing that while $\delta W_{l,\mathrm{ss}}$ and δU_F are reversible, they are not perfect differentials. The heat flow is reversed exactly only if the return path takes us through the same sequence of states as the forward path. If, however, we traverse a loop by modulating two circuit parameters independently, then the contributions from the various sides of the loop do not sum to zero. In the case of δU_F this follows immediately from the fact that δU_F is the generalization of $dU - p\,dV$ for the equilibrium case. In the case of $\delta W_{l,\mathrm{ss}}$ examples can easily be constructed to show that in general

$$\oint \delta W_{l,\mathrm{ss}} \neq 0.$$

The temperature T_N is typically (but not necessarily) higher than the ambient temperature. Thus a fluctuational entropy gain $dS = dQ_F/T_N$ can be less than the entropy taken from the reservoir at a lower temperature. This is, however, not a violation of the second law since these entropy changes take place against a background of steady-state entropy generation.

APPENDIX. DETAILS OF TUNNEL DIODE CIRCUITS

Consider a circuit as shown in Fig. 4. The main point is that the resistor of Fig. 1 has been replaced by a tunnel diode. (We could equally well consider its replacement by any other device which passes one electron at a time.) The stochastic variable is now the charge imbalance between the two capacitors, i.e., the net charge which has been introduced into the junction between the two capacitors. This must change by one electron at a time, through tunneling via either device. We thus have a ladder of possible charge levels, as shown in Fig. 5, and must go up and down this ladder one step at a time. In the steady state if the flow vanishes as we go far along the ladder, this necessitates

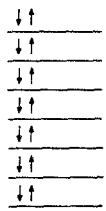

Fig. 5. A ladder of states with transitions allowed only between adjacent levels.

detailed balance, and therefore the analogies to an equilibrium distribution function[3,4] are possible. Furthermore, the analysis of Section 4 of Ref. 8 shows that if the tunnel diode has a large enough area, then the continuum approximation of Eq. (2) becomes accurate.

If we now return to the circuit of Fig. 1, we depart from detailed balance, or at least we lose it in its most literal and rigorous sense. For simplicity consider a heavily forward-biased tunnel diode, for example, near its current minimum. The electronic motion there is almost entirely *in* the direction of net current flow, with very little reverse tunneling from the valence band to the conduction band. Thus the diode current there only discharges the capacitance; the charging current must come through the resistor. The diode discharge current, however, still involves the loss of one electron at a time, whereas in the resistor the elementary stochastic event is the motion of an electron through a mean free path, and typically corresponds to a charge change at the capacitor of much less than one electron. Thus in the steady state the capacitive discharge of one electron at a time is compensated by a trickle of much smaller charge elements through the resistor. Thus rigorous detailed balancing does not exist. It is therefore not necessary to go to two dimensions to find an example of the circulation effects which accompany deviation from detailed balance. Of course, if all the elementary transitions in the system, including the jump of one whole electron at a time, take us between states which are close to each other measured on the scale of the distribution function variation, then we can still invoke the Fokker–Planck equation as a good approximation.

There may be some squeamishness about representing a system with discrete states as typified in Fig. 4 through a continuous distribution function $\rho(q)$. In that case we can consider a closely related system in which one or both of the electron passage devices are placed in parallel with large ordinary resistors. The resistors are chosen to be large enough so as not to influence the current flow appreciably. At the same time they do occasionally permit charge transitions of much less than one electron at a time, and thereby serve to give us a whole series of ladders, all with very similar kinetics, but resulting in a genuinely continuous density distribution $\rho(q)$.

REFERENCES

1. R. Landauer and J. W. F. Woo, *Phys. Rev. A* 6:2204 (1972).
2. H. Haken (ed.), *Synergetics, Cooperative Phenomena in Multi-Component Systems*, B. G. Teubner, Stuttgart (1973).
3. H. Haken, Cooperative effects in systems far from thermal equilibrium and in non-physical systems, *Rev. Mod. Phys.*, to be published.

4. R. Graham, *Statistical Theory of Instabilities in Stationary Nonequilibrium Systems with Applications to Lasers and Nonlinear Physics in Quantum Statistics in Optics and Solid State Physics,* Springer Verlag, Berlin (1973).
5. I. Prigogine, G. Nicolis, and A. Babloyantz, *Phys. Today* **25**(11):23 (1972); I. Prigogine, G. Nicolis, and A. Babloyantz, *Phys. Today* **25**(12):38 (1972).
6. G. Nicolis, in *Advances in Chemical Physics*, Vol. XIX, I. Prigogine and S. A. Rice, eds., Wiley–Interscience, New York (1971).
7. R. L. Stratonovich, *Topics in the Theory of Random Noise*, Gordon and Breach, New York (Vol. I, 1963; Vol. II, 1967).
8. R. Landauer, *J. Appl. Phys.* **33**:2209 (1962).
9. R. Landauer, in *Physik 1971*, B. G. Teubner, Stuttgart (1971), p. 286.
10. R. Landauer and J. W. F. Woo, in *Synergetics, Cooperative Phenomena in Multi-Component Systems*, H. Haken, ed., B. G. Teubner, Stuttgart (1973), p. 97.
11. C. H. Bennett, Logical reversibility of computation, *IBM J. Res. Develop.*, **17**:525 (1973).
12. R. Landauer, *IBM J. Res. Develop.* **5**:183 (1961).
13. N. G. van Kampen, in *Fluctuation Phenomena in Solids*, R. E. Burgess, ed., Academic Press, New York (1965).
14. J. R. Manning, *Diffusion Kinetics for Atoms in Crystals*, Van Nostrand, Princeton, New Jersey (1968), Section 5-1.
15. R. Landauer and J. W. F. Woo, *J. Appl. Phys.* **42**:2301 (1971).
16. R. W. Keyes and R. Landauer, *IBM J. Res. Develop.* **14**:152 (1970).