PHYSICAL REVIEW A VOLUME 9, NUMBER 3 MARCH 1974

New thermodynamic variables for a nonequilibrium system*

Aaron B. Corbet†
Biophysics Research Division, Institute of Science and Technology, University of Michigan, Ann Arbor, Michigan 48105
(Received 26 January 1973; revised manuscript received 7 September 1973)

The information-theoretic formulation of statistical mechanics is applied to an ideal gas transporting energy without transporting mass. It is shown that the thermal state variables fracture into distinct populations, thus exhibiting a radical sort of local nonequilibrium. The partition function is found in exact form, for an arbitrarily large temperature gradient. In an approximation, it is shown that these results reduce to the known ones for equilibrium, and near equilibrium. The same approximation allows us specific equations for the temperature, pressure tensor, and the entropy of the system, in a far-from-equilibrium case, in terms of the new thermal variables of the system. The general entropy functional, and its variation, are discussed.

I. INTRODUCTION

In what follows a statistical-mechanical theory is developed for a specific system arbitrarily far from equilibrium; viz., a stationary ideal gas transporting energy without transporting mass. A few notational conveniences will be freely introduced along the way for the sake of economy.[1]

This system has received a fair amount of study[2-15] by more conventional methods—in the near-equilibrium region—and it happens that the approach of one group of workers is of special relevance to us here. This approach is known as the half-range technique.[5-13] We will have more to say on this in Sec. V, when we treat the near-equilibrium case, and again in Sec. VI, when we compare our final results.[2-22] For the present, it will be worthwhile to comment on the difference between this technique and the population splitting that we will arrive at here.

The present analysis proceeded from the fundamental papers of Jaynes,[23,24] without further reference to the literature, perhaps to some advantage, and is self-contained. Only after completing the study was a literature search made for comparison. There should be no question then of a mere grafting into the present analysis of this prior half-range technique. Rather, the splitting which we will arrive at here has been forced on us purely by logical considerations. In fact, from the theoretical point of view, the *proof* of this splitting is the central result of this paper, and the entire subsequent analysis hinges on it. By comparison, the half-range technique developed by previous investigators has always been introduced purely *ad hoc*, and employed relatively unsystematically. Moreover, with the exception of Krook,[8] it appears to have been conceived entirely as a computational device, and no physical meaning has been assigned to the half-range variables. Lastly, while the treatment of these prior workers is for near-equilibrium, and apart from the introduction of these half-range probability densities, is based on the conventional linear approaches, the present development is completely general and is tied *in no way* to the severely limiting assumptions of a near-equilibrium theory. While we have no intention here of slighting the work of these previous researchers, it seems that these differences must be understood if one is to appreciate the logical cohesion of the present analysis, and the fact that it gives us a perfectly general theory.

II. GENERALITIES

The macroscopic state of an arbitrary system is completely characterized by some set of extensive variables $\{A_l\}$, or equivalently by the conjugate set of intensive variables $\{\lambda_l\}$. In terms of these latter, let us distinguish the following types of nonequilibrium:

Global. (a) Spatial dependence; $\lambda_l \to \lambda_l(\vec{r})$. (b) Time dependence; $\lambda_l \to \lambda_l(t)$.

Local. (c) The appearance, at the local level, of λs other than those conjugate to collisional invariants; e.g., $\lambda_l(\vec{r}, t)$ for $l > 3$ (see below). (d) Partitioning, at the local level, into populations, γ, characterized by different values of the intensive variables; $\lambda_l(\vec{r}, t) \to \lambda_l^\gamma(\vec{r}, t)$.

Although the present section allows for each of these types of nonequilibrium, in what follows we will be mainly concerned with type (d). Type (b) will be treated in a subsequent paper. Type (d) is the most radical of the departures from equilibrium, since it means that thermal exchange (e.g., equipartitioning of energy) is somehow interfered with.

We will assume some familiarity with the information theoretic approach to statistical mechan-

ics,[23,24] so that the following variation on the conventional argument need not be stated in full detail.

It will be more convenient in what follows to work in terms of velocities rather than momenta; the corresponding graining factor for the phase space should then be h/m, and since this will occur frequently in the equations, we will give it a special symbol: $\mathfrak{h}$. Consider a system of N particles, and let $\vec{\mu} \equiv (\vec{v}_1, \ldots, \vec{q}_N)$ denote a point in the microstate space. We are going to assume that we have instruments allowing us to measure the means of certain local observables, $\vec{A}_l(\vec{\mu}, \vec{r})$, for different regions γ of the microstate space. We find the correct probability density, $\sigma_t(\vec{\mu})$, by maximizing

$$S = -k \int \sigma_t(\vec{\mu}) \ln \sigma_t(\vec{\mu})\, d\vec{\mu}/\mathfrak{h}^{3N} N! \tag{1}$$

subject to all the known constraints; that is, subject to

$$\int_\gamma \sigma_t(\vec{\mu}) \vec{A}_l(\vec{\mu}, \vec{r})\, d\vec{\mu}/\mathfrak{h}^{3N} N! = \langle \vec{A}_l \rangle_\gamma(\vec{r}, t) \tag{2}$$

for all observables l, of all regions γ, and points $\vec{r}$. Thus we must add to the integrand of (1) the quantity

$$\sigma_t(\vec{\mu}) \sum_\gamma \sum_l \int \vec{\lambda}_l^\gamma(\vec{r}, t) \cdot \vec{A}_l(\vec{\mu}, \vec{r})\, d\vec{r}, \tag{3}$$

where the $\vec{\lambda}_l^\gamma(\vec{r}, t)$ are the conjugate Lagrange multipliers; we have allowed $\vec{A}_l$ to be a vector, without loss of generality, so as to have a more compact notation. The requirement of normalization can be included by setting $A_0 = 1$. The condition for an extremum of (1) then requires that the piecewise continuous function σ satisfy each of the γ Euler equations[25]

$$\frac{\partial}{\partial \sigma}\left[\sigma \ln \sigma + \sigma\left(\lambda_0 + \sum_l \int \vec{\lambda}_l^\gamma(\vec{r}, t) \cdot \vec{A}_l(\vec{\mu}, \vec{r})\, d\vec{r}\right)\right] = 0, \tag{4}$$

which yields

$$\sigma_t(\vec{\mu}) = \Xi^{-1} \exp\left(-\sum_l \int \vec{\lambda}_l^\gamma(\vec{r}, t) \cdot \vec{A}_l(\vec{\mu}, \vec{r})\, d\vec{r}\right) \tag{5}$$

in each region γ, where

$$\Xi \equiv \exp[\lambda_0 + 1] = \sum_\gamma \int_\gamma \exp\left(-\sum_l \int \vec{\lambda}_l^\gamma(\vec{r}, t) \cdot \vec{A}_l(\vec{\mu}, \vec{r})\, d\vec{r}\right) d\vec{\mu}/\mathfrak{h}^{3N} N! \equiv \sum_\gamma \Xi_\gamma. \tag{6}$$

When this value of σ is substituted in (1), the corresponding quantity is called the (generalized) entropy, and will be denoted by $\mathfrak{S}$.

There are a couple of points which should be made clear straight away. An arbitrary partitioning of the microstate space into the set $\{\gamma\}$ does not actually impose anything on the analysis, since the corresponding $\{\vec{\lambda}_l^\gamma(\vec{r}, t)\}$ will eventually be defined purely in terms of the observables. It may well be that these are such that the $\lambda^\gamma(\vec{r}, t)$ become the same for different regions γ, and we simply end up with a redundant, although correct, description. In equilibrium, for example, all $(\vec{r}, t)$ as well as γ dependence become superfluous, and the analysis reduces to the familiar equilibrium results. *The converse, that nothing is imposed by lumping populations together, is not true, however.* Secondly, different means for different regions of a given partitioning does not necessitate that the conjugate λ^γ's be different. In the case of equilibrium, again, one can observe many different mean energies appropriate to the different regions of an arbitrary partitioning, but the conjugate λ^γ's are all the same. This is perhaps best understood in terms of the formal relation between the observables and the multipliers, which follows.

The observables defined by (2) give the mean contribution of each population γ to the total mean, $\langle \vec{A}_l \rangle(\vec{r}, t)$, as measured over the entire microstate space, and contain a statistical weight for the point $\vec{\mu}$ to fall within the region γ. For some purposes, it may be of advantage to work *within* a given population; i.e., assume that $\vec{\mu}$ is already there, and see what the various means are. For example, the physical meaning of the λs, and of $\mathfrak{S}$, become clearer in these terms. To this end we introduce conditional probabilities[26]; thus, e.g.,

$$\sigma_t(\vec{\mu}) = \sigma_t(\gamma)\sigma_t(\vec{\mu}|\gamma), \tag{7}$$

where $\sigma_t(\gamma)$ is the probability that the system microstate is in the region γ, and $\sigma_t(\vec{\mu}|\gamma)$ is the conditional probability for the occurrence of the chosen microstate, given that we are already in the region γ. Then, denoting means within a population by the subscript $\subset\gamma$, we have

$$\langle \vec{A}_l \rangle_\gamma(\vec{r}, t) = \sigma_t(\gamma) \int_\gamma \sigma_t(\vec{\mu}|\gamma) \vec{A}_l(\vec{\mu}, \vec{r})\, d\vec{\mu}/\mathfrak{h}^{3N} N! = \sigma_t(\gamma) \langle \vec{A}_l \rangle_{\subset\gamma}, \tag{8}$$

where

$$\sigma_t(\gamma) = \int_\gamma \sigma_t(\vec{\mu})\, d\vec{\mu}/\mathfrak{h}^{3N} N! = \Xi_\gamma/\Xi = \langle n \rangle_\gamma/N. \tag{9}$$

Thus, in terms of the mean contribution of the

populations, we have

$$\langle \vec{A}_l \rangle_\gamma(\vec{r}, t) = -\delta \ln \Xi / \delta \vec{\lambda}_l^\gamma \tag{10}$$

and in terms of means within populations

$$\langle \vec{A}_l \rangle_{\subset\gamma}(\vec{r}, t) = -\delta \ln \Xi_\gamma / \delta \vec{\lambda}_l^\gamma, \tag{11}$$

where the right-hand side is the functional derivative, as defined, e.g., in Gelfand and Fomin.[25]

At this point we can make the following useful definitions.

Definition 1. We will call a scalar-valued function, $A(\vec{\mu}, \vec{r})$, a *generic variable* when it is defined in such a way as to be independent of the particular subspace serving as its domain.

Example (i). $A = \frac{1}{2} m v_{\nu i}{}^2$ is a generic variable. Its mean on the subspace $\gamma = \{v_{7x}\}$ is $\langle \frac{1}{2} m v_{7x}^2 \rangle_\gamma(\vec{r})$. Its mean on the entire microstate space is

$$\left\langle \tfrac{1}{2} m \sum_{\nu=1}^{N} \sum_{i=1}^{3} v_{\nu i}^2 \right\rangle (\vec{r}).$$

Example (ii). $A = (1/q_{\nu y}) \sin(q_{\nu i}^2 q_{\nu j})$ is a generic variable. Its mean on the subspace $\gamma = \{q_{1x}, q_{1z} = 1 - q_{1x}\}$ is $(1/q_{1y})\langle \sin(q_{1z}^2 q_{1x}) \rangle_\gamma(\vec{r})$. Its mean over the entire microstate space is

$$\left\langle \sum_{\nu} (1/q_{\nu y}) \sum_{i \neq j} \sin(q_{\nu i}^2 q_{\nu j}) \right\rangle (\vec{r}).$$

The last example illustrates that some ambiguity will arise unless we orient the space so as to give a definite meaning to expressions with multiple subscripts; e.g., does $q_{\nu i}^2 q_{\nu j}$ on the above γ mean $q_{1z}^2 q_{1x}$ or $q_{1x}^2 q_{1z}$? We will limit the definition of generic variables to those that take the coordinates in a definite order; viz., as cyclic permutations of x, y, z. Thus, for the above γ, $q_{\nu i}^2 q_{\nu j}$ can only be $q_{1z}^2 q_{1x}$; $q_{\nu i} q_{\nu j}^2$, on the other hand, would be $q_{1x}^2 q_{1z}$. The same remark applies to particle order: we arbitrarily pick a given order, and permute particles cyclically. (This will be irrelevant for identical particles.)

Definition 2. By the *intensive state variables* of a system, we mean all the distinct nonvanishing Lagrange multipliers conjugate to means of generic variables over all partitionings $\{\gamma\}$. (Owing to the partitioning, these will not all be independent, in general.)

Definition 3. By the *extensive state variables* of a system, we mean the observables conjugate to the intensive state variables.

In the case of equilibrium, the system is generally completely isotropic, and so the state variables, as well as being global, correspond to means over the entire microstate space. We cannot expect this to be true, in general, for nonequilibrium systems.

These definitions, while logically complete, still do not furnish us with criteria for *finding* these state variables. This is, in fact, an omission in all the information-theoretic discussions we have seen. No real distinction is drawn between observables, independent observables, and (thermal) state variables. For example, it would seem possible to add on to, or replace, the term $\lambda_3 v^2$ in the equilibrium distribution with $\lambda_5 v^4$ or $\lambda_7 v^6$ since $\langle v^4 \rangle$ and $\langle v^6 \rangle$ are also observables. Yet we know this does not give the correct equilibrium distribution. In order to eliminate this ambiguity, we propose the following criteria.

Criterion of sufficient information. The extensive state variables correspond to generic variables whose nonvanishing net means provide a maximal set of mutually independent observables. (These may be infinite in number, generally.)

Criterion of simplicity. Among competing candidates for state variables, nature chooses that set which, by some clear standard, is simplest.

We henceforth assume that (5) satisfies these criteria.

Substituting (5) in (1) gives

$$\mathfrak{S} = k \left(\ln \Xi + \sum_\gamma \sum_l \int \vec{\lambda}_l^\gamma \cdot \langle \vec{A}_l \rangle_\gamma(\vec{r}, t)\, d\vec{r} \right), \tag{12}$$

which, by (6) and (10) gives the entropy entirely in terms of the intensive state variables; i.e., as the functional $\mathfrak{S}[\{\vec{\lambda}_l^\gamma(\vec{r}, t)\}]$. If we now use (8) and (9) in (12), we get

$$\begin{aligned} \mathfrak{S} &= k \left(\sum_\gamma \sigma_t(\gamma) \left[\ln \Xi_\gamma + \sum_l \int \vec{\lambda}_l^\gamma \cdot \langle \vec{A}_l \rangle_{\subset\gamma}(\vec{r}, t)\, d\vec{r} \right] \right. \\ &\quad \left. + \sum_\gamma \sigma_t(\gamma) \ln \sigma_t(\gamma) \right) \\ &\equiv \sum_\gamma \sigma_t(\gamma) \mathfrak{S}(\vec{\mu} | \gamma) + \Delta_\gamma \mathfrak{S}, \end{aligned} \tag{13}$$

where we have made the obvious identifications. We see that the entropy decomposes in a natural way into a weighted mean of the entropies *within* the given populations, plus an "entropy of mixing" term.

It is of interest to see how the various dependencies affect the magnitude of $\mathfrak{S}$. One way of doing this is to compare between systems with different constraints in common; i.e., we vary $\mathfrak{S}$, subject to these common constraints, to find the curve $\{\langle \vec{A}_l \rangle_\gamma(\vec{r}, t)\}$ which maximizes $\mathfrak{S}$. (Whereas previously we found $\mathfrak{S}$ by finding the curve $\sigma_t(\vec{\mu})$ which maximized S.) Let us suppose, then, that we are comparing between systems which share some common set of *net* local observables $\langle \vec{A}_{l'} \rangle(\vec{r}, t) = \sum_\gamma \langle \vec{A}_{l'} \rangle_\gamma(\vec{r}, t)$. Let the Lagrange multipliers be $\vec{g}_{l'}(\vec{r}, t)$. Then the condition of an extremum requires that each of the

$l+\gamma$ derivatives vanish:

$$\frac{\delta}{\delta\langle\vec{A}_l\rangle_\gamma}\left(\frac{1}{k}\mathfrak{S}[\{\langle\vec{A}_l\rangle_\gamma(\vec{r},t)\}]+\sum_{l'}\int\vec{g}_{l'}(\vec{r})\cdot\sum_\gamma\langle\vec{A}_{l'}\rangle_\gamma(\vec{r},t)\,d\vec{r}\right)=0. \tag{14}$$

The functional derivatives of (12) are readily seen to be

$$\frac{1}{k}\frac{\delta\mathfrak{S}}{\delta\langle\vec{A}_l\rangle_\gamma}=\vec{\lambda}_l^\gamma(\vec{r},t), \tag{15}$$

whence (14) becomes

$$\vec{\lambda}_l^\gamma(\vec{r},t)+\vec{g}_{l'}(\vec{r},t)\delta_{ll'}=0. \tag{16}$$

This means that among all those systems sharing the above common set of local constraints (net local observables), one system has the maximum entropy for which all the other intensive variables, $l\neq l'$, vanish, and for which the microstate partitioning is superfluous for the remaining variables. Which is what we should expect. Similarly, one can compare between systems with a common set of net global, time-independent, observables $\langle\vec{A}_{l'}\rangle^*$. The variational derivatives then become

$$\vec{\lambda}_l^\gamma(\vec{r},t)+\vec{G}_{l'}\delta_{ll'}=0, \tag{17}$$

where the $\vec{G}_{l'}$ are the constant multipliers conjugate to the global observables. This means, then, that among those systems having these observables in common, that system has the maximum entropy for which the intensive variables conjugate to all other observables, $l\neq l'$, vanish, all microstate partitioning is superfluous, and in addition, all space and time dependence of the remaining λs vanish. Which, again, is what we should expect. By introducing more complex conditions to be held in common, via the Lagrange multiplier method, one can get a variety of conditions on the λs; inversion of the relations (10) then yields the maximizing curve $\{\langle\vec{A}_l\rangle_\gamma(\vec{r},t)\}$. However, such inversion may often be difficult to accomplish in practice.

There is one final, general, thermodynamic property which we will note here, since this will be a central result for what follows. If ever we have equations of state of the form

$$F\langle A_i\rangle_\gamma(\vec{r},t)=\delta(G\langle A_j\rangle_\gamma)/\delta\lambda_k^\gamma \tag{18}$$

true for all γ, where the multiplicands F and G are arbitrary nonzero functionals, then by summing both sides over γ, we see that in the case that the net observable $\langle A_j\rangle(\vec{r},t)$ vanishes while the net observable $\langle A_i\rangle(\vec{r},t)$ does not, it clearly must be the case that

$$\lambda_k^\gamma(\vec{r},t)=0 \text{ for all } \gamma$$

or

$$\lambda_k^\gamma(\vec{r},t)\neq\lambda_k^{\gamma'}(\vec{r},t) \text{ for some } \gamma\neq\gamma'. \tag{19}$$

This is actually true whether or not the $\langle A_l\rangle_\gamma$s are themselves state variables.[26a]

III. STATIONARY CLASSICAL IDEAL GAS

In this case, the only observables possible are functions of the particle velocities, and the intensive variables of the system can only depend on the particles *at* the chosen point. (Whereas, in general, one could measure a field, say in a plasma, which would be a function of all the particle positions, and could be at a point in the system where the particle density was zero.) We will define a general moment vector, and assume that some subset of these moments can always be found as the state variables of the system. Let

$$\vec{V}^l\equiv\sum_{\nu=1}^{N}\vec{v}_\nu^{\,l}, \tag{20}$$

where the sum is over particles, and the components of $\vec{v}^l$ are all the distinct products of the form

$$v_{a_1}v_{a_2}\cdots v_{a_l},\quad a_i=x,y,z. \tag{21}$$

Each $\vec{v}^l$ will thus have $(l+2)!/l!2!$ components. The preceding comments can then be summarized formally as

$$\vec{\lambda}_l^\gamma(\vec{r},t)\cdot\vec{A}_l(\vec{\mu},\vec{r})-\vec{\lambda}_l^\gamma(\vec{r})\cdot\sum_\nu\vec{v}_\nu^{\,l-1}\delta(\vec{r}-\vec{q}_\nu). \tag{22}$$

The exponent of (5) thus becomes

$$-\sum_l\sum_\nu\vec{\lambda}_l^\gamma(\vec{q}_\nu)\cdot v_\nu^{l-1}. \tag{23}$$

The independence of the particles then allows (5) to factor into a product of one-particle probability densities as follows:

$$\sigma(\vec{\mu})=N!\prod_\nu\sigma_1(\vec{v}_\nu,\vec{q}_\nu), \tag{24}$$

where

$$\sigma_1(\vec{v},\vec{q}=\vec{r})=\xi^{-1}\exp\left(-\sum_l\vec{\lambda}_l^\gamma(\vec{r})\cdot\vec{v}^{l-1}\right),$$
$$\xi\equiv\int\exp\left(-\sum_l\vec{\lambda}_l^\gamma(\vec{r})\cdot\vec{v}^{l-1}\right)d\vec{v}\,d\vec{r}/\hbar^3. \tag{25}$$

We have made use here of the fact that the equivalence of the particles causes $\{\gamma\}$ to partition each subspace $\{\vec{v}_\nu,\vec{q}_\nu\}$ in exactly the same way. The independence of the particles then means that

$\vec{\lambda}_l^{\gamma}(\vec{q}_\nu)$ in (23) depends on position only through the one-particle coordinate, which is already given. Thus γ, in (25), depends on the one-particle velocity only.

A local, one-particle partition function can be defined by

$$\zeta(\vec{r}) \equiv \xi \int \sigma_1(\vec{v}, \vec{r})\, d\vec{v}/\mathfrak{h}^3, \tag{26}$$

so that $\xi = \int \zeta(\vec{r})\, d\vec{r}$. Since $\Xi = \xi^N/N!$, use of (24) and (26), in (10) and (11) respectively, yields

$$\langle \vec{\mathfrak{V}}^{l-1} \rangle_\gamma(\vec{r}) = -N\partial \ln\zeta(\vec{r})/\partial\vec{\lambda}_l^{\gamma} = \langle n \rangle_\gamma(\vec{r}) \langle \vec{\mathfrak{V}}^{l-1} \rangle_{\subset\gamma}(\vec{r}), \tag{27}$$

$$\langle \vec{\mathfrak{V}}^{l-1} \rangle_{\subset\gamma}(\vec{r}) = -\partial \ln\zeta_\gamma(\vec{r})/\partial\vec{\lambda}_l^{\gamma}, \tag{28}$$

where $\zeta_\gamma(\vec{r}) \equiv \xi \int_\gamma \sigma_1(\vec{v}, \vec{r})\, d\vec{v}/\mathfrak{h}^3$, and $\xi_\gamma \equiv \int \zeta_\gamma(\vec{r})\, d\vec{r}$. There is one further set of relations which will be useful in the following. Analogous to (7), we can define a local probability density, $\sigma_1(\vec{v}|\vec{r})$, by

$$\sigma_1(\vec{v}, \vec{r}) = \sigma_1(\vec{r})\, \sigma_1(\vec{v}|\vec{r}) \tag{29}$$

so that, explicitly,

$$\begin{aligned} &\sigma_1(\vec{r}) = \zeta(\vec{r})/\xi, \\ &\sigma_1(\vec{v}|\vec{r}) = \zeta^{-1}(\vec{r}) \exp\left(-\sum_l \vec{\lambda}_l^{\gamma}(\vec{r}) \cdot \vec{\mathfrak{V}}^{l-1}\right). \end{aligned} \tag{30}$$

We are now in a position to obtain an expression for the local entropy. Writing $\mathfrak{S} = \int \mathfrak{s}(\vec{r})\, d\vec{r}$, we get, by substitution of (24) and (29) in (1),

$$\mathfrak{s}(\vec{r}) = \Delta_{\langle n \rangle(\vec{r})} \mathfrak{s} + \langle n \rangle(\vec{r})\, \mathfrak{s}_1(\vec{v}|\vec{r}), \tag{31}$$

where

$$\begin{aligned} &\langle n \rangle(\vec{r}) = N\sigma_1(\vec{r}), \\ &\Delta_{\langle n \rangle(\vec{r})} \mathfrak{s} = -k\langle n \rangle(\vec{r}) \ln[\langle n \rangle(\vec{r})/e], \\ &\mathfrak{s}_1(\vec{v}|\vec{r}) = -k \int \sigma_1(\vec{v}|\vec{r}) \ln\sigma_1(\vec{v}|\vec{r})\, d\vec{v}/\mathfrak{h}^3. \end{aligned} \tag{32}$$

Alternatively, if we now use (7), respecting the preceding remarks on $\{\gamma\}$, we can get the finer decomposition

$$\mathfrak{s}(\vec{r}) = \Delta_{\langle n \rangle_\gamma(\vec{r})} \mathfrak{s} + \sum_\gamma \langle n \rangle_\gamma(\vec{r})\, \mathfrak{s}_1(\vec{v}|\vec{r}, \gamma), \tag{33}$$

where

$$\begin{aligned} &\Delta_{\langle n \rangle_\gamma(\vec{r})} \mathfrak{s} = -k \sum_\gamma \langle n \rangle_\gamma(\vec{r}) \ln[\langle n \rangle_\gamma(\vec{r})/e], \\ &\mathfrak{s}_1(\vec{v}|\vec{r}, \gamma) = -k \int_\gamma \sigma_1(\vec{v}|\vec{r}, \gamma) \ln\sigma_1(\vec{v}|\vec{r}, \gamma)\, d\vec{v}/\mathfrak{h}^3. \end{aligned}$$

IV. ONE-DIMENSIONAL HEAT FLOW

We are going to present a very specific, somewhat idealized, model, and complete the analysis in exact form for that model. Our results will turn out to have greater generality than this, as will be discussed.

We will let the ideal gas of N particles be confined in a cylindrical volume, $2\pi RL$, between two flat, parallel walls, held at the temperatures T_1 and T_2, respectively. The nature of the boundary is important. By "flat" walls, we mean ones for which the vibration of the wall molecules is in the axial, z direction only. The cylindrical wall, on the other hand, is perfectly adiabatic. If we wish, we can interpret this as meaning that the molecules of which it is composed are at 0°K and have infinite inertia, so that the gas particles undergo perfectly reflecting collisions. In an actual experimental situation, if $R >> L$, and the end walls are constructed of clean, crystalline, surfaces, relatively free of imperfections, the model will be well approximated in the interior of the cylinder.

With this specific situation we are now prepared to reduce (25) to a finite form. For a given total number of particles, N, there are three external constraints on the system: the volume, and the two temperatures T_1 and T_2. We choose to write these external constraints as

$$\begin{aligned} &A = \text{fixed volume}, \\ &B = \text{fixed } T_1 + T_2, \\ &C = \text{fixed } T_1 - T_2. \end{aligned} \tag{34}$$

These, in turn, produce the following *internal* constraints:

$$A \Rightarrow m\langle n \rangle(z), \quad \text{mass density} \tag{35}$$

$$B \Rightarrow \langle u \rangle(z), \quad \text{energy density} \tag{36}$$

$$C \Rightarrow \langle h_z \rangle(z), \quad \text{energy flux.} \tag{37}$$

The implication here indicates a *necessary* condition. This does not mean that the effects of (34) are separable; e.g., the temperature gradient $T_1 - T_2$ will clearly affect the spatial density, etc. However, the independence of the constraints (34) necessitates the independence of the net observables (35)–(37). These observables are the net means of the generic variables $m\sum_\nu \delta(\vec{r} - \vec{q}_\nu)$, $\frac{1}{2}m\sum_\nu \delta(\vec{r} - \vec{q}_\nu) v_{\nu i}^2$, and $\frac{1}{2}m\sum_\nu \delta(\vec{r} - \vec{q}_\nu) v_{\nu i}^2 v_{\nu z}$, respectively. Clearly, this is a maximal set, since there can be no more independent thermal variables than independent external constraints. Thus we have satisfied the criterion of sufficient information of Sec. II, here. Moreover, these are the lowest nonvanishing moments, as we will see directly. In accordance with the second criterion of Sec. II, we can therefore expect the state variables to derive from this set, possibly together with the other low moment, $m\sum_\nu \delta(\vec{r} - \vec{q}_\nu) v_{\nu i}$; i.e., for an appropriate partitioning $\{\gamma\}$, the state variables should be drawn from the set of means

over subspaces γ of these four generic variables, *and no others*.

The further condition of stationarity implies that net mass flux

$$m\langle\vec{V}\rangle(\vec{r})=0 \tag{38}$$

and that there be no gradients in the direction of any fluxes.[27] The requirement of flat walls means that there is no correlation between v_x, v_y, v_z, as such a correlation would be generated at the walls if at all, and conservation of momentum rules out that possibility. Thus the variables are separable. The symmetry of the system, separability of variables, and the condition (38), then lead to the vanishing of all second moments except $\langle V_x^2\rangle=\langle V_y^2\rangle$, and $\langle V_z^2\rangle$. Similarly, the only nonvanishing third moment is $\langle V_z^3\rangle=\langle h_z\rangle=\langle\vec{h}\rangle$.

We now make the following partitioning: Guided by the symmetry of the system, we divide up the velocity space into the regions $\{v_z\geq 0\}$, $\{v_z\leq 0\}$, and $\{v_r=(v_x^2+v_y^2)^{1/2}\}$. We then have the following observables, derived from (35)–(38):

$$m[\langle n\rangle_+(z)+\langle n\rangle_-(z)]\equiv\rho_m(z), \tag{39}$$

$$m[\langle V_z\rangle_+(z)+\langle V_z\rangle_-(z)]=0, \tag{40}$$

$$\tfrac{1}{2}m[\langle V_z^2\rangle_+(z)+\langle V_z^2\rangle_-(z)]=E_z,\ \text{const}, \tag{41}$$

$$\tfrac{1}{2}m\langle V_r^2\rangle(z)=E_r(z), \tag{42}$$

$$\tfrac{1}{2}m[\langle V_z^3\rangle_+(z)+\langle V_z^3\rangle_-(z)]\equiv\phi_E,\ \text{const}. \tag{43}$$

This is a total of nine scalar observables, for which we must introduce the nine conjugate Lagrange multipliers $\lambda_1^\pm(z)$, $\lambda_2^\pm(z)$, $\lambda_3^\pm(z)$, $\lambda_3^r(z)$, and $\lambda_4^\pm(z)$. Of these, some should vanish, and we do not expect, owing to (34), that any more than three will be independent. In particular, we should expect that λ_3^r will be a function of λ_3^+ and λ_3^-. It also seems likely that the system will make do with only one of the pairs $\lambda_2^\pm$, or $\lambda_4^\pm$.

The requirement of continuity of σ means that $\lambda_1^+(z)=\lambda_1^-(z)$. We conclude that the one-particle probability density must have the following form, then:

$$\sigma_1(\vec{v},\vec{r})=\xi^{-1}\exp[-(\lambda_1(z)+\lambda_2^\pm(z)v_z+\lambda_3^\pm(z)v_z^2+\lambda_3^r(z)v_r^2+\lambda_4^\pm(z)v_z^3)],\quad v_z\gtrless 0. \tag{44}$$

By substitution in (26)–(28), and (33), the entire thermodynamics of the system are exactly solved—formally, and in principle. There is, in the first place, a computational problem in that we are unable to evaluate (26) in closed form. A little manipulation gives

$$\zeta^\pm(r)=\frac{\pi h^3\exp[-\lambda_1+(\lambda_2^\pm)^2/4\lambda_3^\pm]}{m^3\lambda_3^r}\sum_{n=0}^{\infty}\sum_{k=0}^{2n}\frac{(-1)^n(2n)!\,\Gamma[\frac{1}{3}(2n-k+1)]}{3(2n-k)!\,k!\,n!\,2^k}\frac{(\pm\lambda_2^\pm)^k(\lambda_3^\pm)^{n-k}}{(\pm\lambda_4^\pm)^{(2n-k+1)/3}}, \tag{45}$$

which converges slowly in the general case, so that any equations of state derived through (27) and (28) become quite complicated. Furthermore, for any such equations to be of practical value, we must determine the relations between the λs, and their spatial dependence, which means it might be simpler to do this first where possible. As we will see, it is still possible to draw some important qualitative conclusions.

Proof of nonequilibrium of type (d). The observables $\langle V_z^l\rangle_\gamma$ satisfy relations of the form $\delta(\xi\langle V_z^{l'}\rangle_\gamma)/\delta\lambda_{l+1}^\gamma=\xi\langle V_z^{l''}\rangle_\gamma$, for any $l''=l'+l$; $l,l',l''>1$. Then, owing to (38), if we put λ_k^γ successively equal to $\lambda_2^\pm$, $\lambda_3^\pm$, and $\lambda_4^\pm$, in (18) we obtain

$$\delta(\xi\langle V_z\rangle_+)/\delta\lambda_2^+ +\delta(\xi\langle V_z\rangle_-)/\delta\lambda_2^- =\xi\langle V_z^2\rangle>0\Rightarrow\lambda_2^+\neq\lambda_2^-,$$

or both identically 0, (46)

$$\delta(\xi\langle V_z\rangle_+)/\delta\lambda_3^+ +\delta(\xi\langle V_z\rangle_-)/\delta\lambda_3^- =\xi\langle V_z^3\rangle>0\Rightarrow\lambda_3^+\neq\lambda_3^-,$$

or both identically 0, (47)

$$\delta(\xi\langle V_z\rangle_+)/\delta\lambda_4^+ +\delta(\xi\langle V_z\rangle_-)/\delta\lambda_4^- =\xi\langle V_z^4\rangle>0\Rightarrow\lambda_4^+\neq\lambda_4^-,$$

or both identically 0, Q.E.D. (48)

We thus see that the requirement of no mass flux is a severe restraint on the system, causing all the intensive variables conjugate to the v_z observables to fracture into two distinct populations.

Normalization further requires $\lambda_4^+\geq 0\geq\lambda_4^-$. We will assume $\lambda_3^\pm$ must be positive, as in equilibrium. The positions of the maxima in (44) then depend on $\lambda_2^\pm$; e.g., if $\lambda_2^+>0>\lambda_2^-$, there will be a single maximum (cusp) at $v_z=0$. If $\lambda_2^+<0<\lambda_2^-$, the distribution will be bimodal. Now, experiments in scattering of molecular beams by solid surfaces[16,17] indicate that for a wide variety of incident-beam velocity distributions, and surface conditions, the scattered beam is peaked in the forward v_z direction before it undergoes further intermolecular collisions. This appears to be because there is a higher frequency of collisions with the wall molecules when they are moving out than when they are moving in, so the *effective* velocity of

the wall molecules is outward.[18-20] Thus we would expect forward peaks in the half-distributions leaving the walls, in our system, at least through the Knudsen layer.[2-4] Beyond there, we should at least consider the possibility that these peaks are annihilated through collisions, although the fracturing described by (46)–(48) is propagated. One can satisfy the requirements (39)–(43) with a unimodal distribution, so this case cannot be totally ruled out. In the following, however, we will only consider the bimodal distribution, an example of which is schematized in Fig. 1, for the case of large heat flux. We do so for the following reasons: (i) in the free-molecule limit, such a distribution should hold; (ii) at higher densities, in the ideal gas model, the effect of collisions can well be imagined to be predominantly an exchange of v_z momenta, so that the population differences persist; (iii) in the case of the approximation which we will introduce directly, one can prove fairly rigorously that if σ_1 is continuous it must be bimodal.[28]

The requirements on the λs can then be summarized

$$\lambda_1^+(z)=\lambda_1^-(z), \quad \lambda_2^+(z)\leq 0\leq\lambda_2^-(z),$$
$$\lambda_3^\pm(z)>0, \quad \lambda_4^+(z)\geq 0\geq\lambda_4^-(z). \tag{49}$$

V. SECOND-MOMENT APPROXIMATION

Our progress is greatly aided if we approximate (44) by the union of two Gaussian distributions, as illustrated in Fig. 2. This then leads[28] to the further requirement

$$0<\lambda_3^+(z)<\lambda_3^-(z). \tag{50}$$

The local partition function, (26), now becomes

$$\zeta^\pm(\vec{r})=[\pi^{3/2}m^3/2h^3\lambda_3^r(\lambda_3^\pm)^{1/2}][1+\mathrm{erf}(\alpha^\pm)]$$
$$\times\exp[-\lambda_1+(\alpha^\pm)^2], \qquad \zeta=\zeta^++\zeta^-, \tag{51}$$

where $\mathrm{erf}(\alpha)=2\pi^{-1/2}\int_0^\alpha e^{-t^2}dt$, and we have defined $\alpha^\pm\equiv\mp\lambda_2^\pm/2(\lambda_3^\pm)^{1/2}$, which is positive, by (49).

In fact, if the observables $\langle n\rangle_\pm$, $\langle V_z\rangle_\pm$, $\langle V_z^2\rangle_\pm$ are independent on the half-spaces $v_z\gtrless 0$, respectively, we should have, according to the criterion of simplicity of Sec. II, that better than an approximation, this *is* the partition function. We frankly suspect this is so. It would seem that if one set of these variables, say for $v_z>0$, is dependent, then so should the other set be. But then so would the *net* observables be, and we know this is not so. We will not argue more strongly for these being the state variables of the system, at this point.

We now consider three limiting cases.

Equilibrium

In this case, the third external constraint of (34) vanishes, (37) vanishes, $\alpha^\pm$ of (51) vanishes, and all spatial dependence vanishes. Then $\sigma_1(\vec{r})=\zeta(\vec{r})/\xi$ is just $1/V$, where V is the volume here; $\lambda_2^\pm=0$; $\lambda_3^\pm=\lambda_3^r=m/2kT$; and $\langle v_z^2\rangle_{\subset\pm}=\langle v_z^2\rangle=\frac{1}{2}\langle v_r^2\rangle$. Substitution in (31), or the equivalent formulas, then yields the well known alternative forms for the entropy

$$\mathfrak{S}=\mathfrak{N}k\ln[(e^{5/2}V/\mathfrak{N})(2m\pi kT/h^2)^{3/2}], \tag{52}$$

$$\mathfrak{S}=\mathfrak{N}k\ln[e^{5/2}(4m\pi/3h^2)^{3/2}V\mathfrak{U}^{3/2}\mathfrak{N}^{-5/2}], \tag{53}$$

where we have defined the state variables $\mathfrak{N}\equiv\langle n\rangle^*=N$, and $\mathfrak{U}\equiv\frac{1}{2}m\mathfrak{N}\langle v^2\rangle^*$.

Near equilibrium

Everything can be gotten from (51) by putting $\mathrm{erf}(\alpha)\simeq 2\pi^{-1/2}\alpha$. Instead, we will adopt a perturbation approach here, and calculate the first-order quantities, denoted by superscript (1), in terms of the equilibrium quantities, denoted by superscript (0). Thus

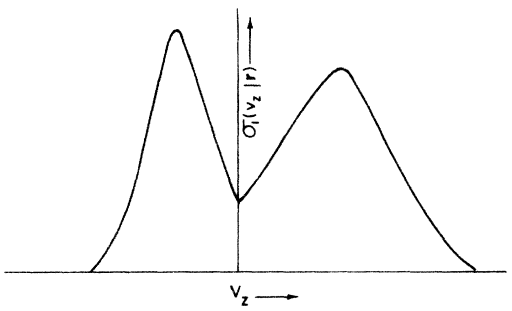

FIG. 1. Example of a bimodal velocity distribution, given by (44), for a case of high heat flux.

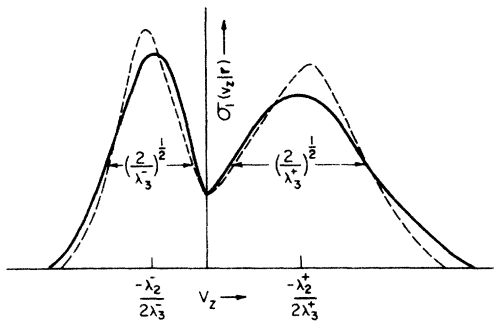

FIG. 2. Second-moment approximation to the distribution of Fig. 1.

$$\sigma_1^{(1)}(\vec{v},\vec{r}) = (\xi^{(1)})^{-1}\exp[-d\lambda_1 - v_z\,d\lambda_2^{\pm} - v_z^2\,d\lambda_3^{\pm} - v_r^2\,d\lambda_3^{r} - \lambda_1^{(0)} - \lambda_3^{(0)}\vec{v}^2]$$
$$\simeq (\xi^{(0)}/\xi^{(1)})[1 - d\lambda_1 - v_z\,d\lambda_2^{\pm} - v_z^2\,d\lambda_3^{\pm} - v_r^2\,d\lambda_3^{r}]\,\sigma_1^{(0)}(\vec{v},\vec{r}). \quad (54)$$

We note that if we had used the original distribution, (44), the expansion analogous to (54) would be formally identical, owing to the vanishing of all mixed moments, to a Grad's thirteen-moment expression[27] for each half-space probability density. The theoretical basis would be quite different, however. Whereas the set of moments for the Hermite tensor expansion is infinite, and only the complete set can, in principle, provide the same information as σ_1, in our case the *only* moments that are necessary to consider are (39)–(43). The information-theoretic approach, supplemented with the two criteria of Sec. II, eliminates all other moments in one blow. Higher moments will appear only artificially, as a result of continuing the *linear* expansion of (44) to get more terms. The meaning of the coefficients in this expansion are, moreover, quite different than in the Hermite tensor expansion. It is perhaps also worth reminding the reader that whereas consideration of two separate probability densities, $\sigma_1(v_z \gtrless 0)$, would appear as a relatively arbitrary element in the standard theory via the Boltzmann equation, this splitting appears here as a natural and logical result of the information theoretic analysis.

From our present expression (54), the local means within populations, $\langle v_z^{\,l}\rangle_{\subset\pm}(\vec{r})$, are readily found to be

$$\langle v_z^{\,l}\rangle_{\subset\pm}^{(1)} = \langle v_z^{\,l}\rangle_{\subset\pm}^{(0)} - C_{l1}^{(0)\pm}\,d\lambda_2^{\pm} - C_{l2}^{(0)\pm}\,d\lambda_3^{\pm} + O(d^2), \quad (55)$$

where

$$C_{lj}^{(0)\pm} = C_{jl}^{(0)\pm} \equiv \langle v_z^{\,j+l}\rangle_{\subset\pm}^{(0)} - \langle v_z^{\,j}\rangle_{\subset\pm}^{(0)}\langle v_z^{\,l}\rangle_{\subset\pm}^{(0)} = \langle v_z^{\,j+l}\rangle_{\subset\pm}^{(0)}\left(1 - \frac{\Gamma[\frac{1}{2}(j+1)]\,\Gamma[\frac{1}{2}(l+1)]}{\pi^{1/2}\Gamma[\frac{1}{2}(j+l+1)]}\right) \quad (56)$$

and we have simply made use of (29) and (30). The net, N-particle local means are given by

$$\langle V_z^{\,l}\rangle^{(1)} = (\xi^{(0)}/\xi^{(1)})\langle n\rangle^{(0)}[\langle v_z^{\,l}\rangle_{\subset+}^{(0)}(1+(-1)^l)(1 - d\lambda_1 - \langle v_r^2\rangle_{+}^{(0)}\,d\lambda_3^{r}) - \langle v_z^{\,l+1}\rangle_{\subset+}^{(0)}(d\lambda_2^{+} + (-1)^{l+1}\,d\lambda_2^{-}) - \langle v_z^{\,l+2}\rangle_{\subset+}^{(0)}(d\lambda_3^{+} + (-1)^{l+2}\,d\lambda_3^{-})], \quad (57)$$

since $\langle v_z^{\,l}\rangle_{\subset\pm}^{(0)} = (\pm 1)^l\langle v_z^{\,l}\rangle_{\subset+}^{(0)}$. If we then define $\beta \equiv 2\xi^{(0)}/\xi^{(1)}$, $\Delta d\lambda \equiv \frac{1}{2}(d\lambda^{+} - d\lambda^{-})$, and $d\bar{\lambda} \equiv \frac{1}{2}(d\lambda^{+} + d\lambda^{-})$, we can write (40) and (43) as

$$m\langle V_z\rangle^{(1)} = m\langle n\rangle^{(0)}\beta[-d\bar{\lambda}_2\langle v_z^2\rangle_{\subset+}^{(0)} - \Delta d\lambda_3\langle v_z^3\rangle_{\subset+}^{(0)}] = 0, \quad (58)$$

$$\tfrac{1}{2}m\langle V_z^3\rangle^{(1)} = \tfrac{1}{2}m\langle n\rangle^{(0)}\beta[-d\bar{\lambda}_2\langle v_z^4\rangle_{\subset+}^{(0)} - \Delta d\lambda_3\langle v_z^5\rangle_{\subset+}^{(0)}] = \phi_E. \quad (59)$$

Substituting (58) in (59) gives

$$\phi_E = c_1\beta m\langle n\rangle^{(0)}\langle v_z^5\rangle_{\subset+}^{(0)}\Delta d\lambda_3, \quad (60)$$

where c_1 is a number of order unity. On the other hand, defining the experimental temperatures $T_{(1)}^{\pm}$ as $(1/3k)m\langle v^2\rangle_{\subset+}^{(1)}$, we find by substitution of (58) in (55),

$$\Delta d\lambda_3 = c_2 m(T_{(1)}^{+} - T_{(1)}^{-})/kT_{(0)}^2, \quad (61)$$

where c_2 is also a number of the order of unity. Now, we argue that the separation between $T_{(1)}^{\pm}$, at a given point, should increase with the local temperature gradient, and also with the mean free path; for small gradients, we can approximate this as a proportionality. Thus, substituting (61) in (60),

$$\phi_E \propto \langle n\rangle^{(0)}\,T_{(0)}^{1/2}\,l\,\partial\bar{T}_{(1)}/\partial z, \quad (62)$$

where l here denotes the mean free path. Thus we arrive at the standard result of kinetic theory.[27]

We observe that, since ϕ_E is independent of z, (60) says that the difference between $d\lambda_3^{+}$ and $d\lambda_3^{-}$ is held constant at all distances between the walls. By (58), we see that $d\bar{\lambda}_2$ is also independent of distance, then, and by (49) this is the difference of two positive quantities characterizing the populations. In this sense, the splitting of the intensive variables is quantitatively preserved through collisions. This is borne out by calculations of these variables in terms of collision integrals.[6]

A more detailed treatment of the near-equilibrium case will be given in a later paper devoted to the spatial dependencies of the thermal variables of this system. Further comparison of the present theory with prior work is given in Sec. VI.

Far from equilibrium; synchronized heat flux

We consider the limiting case of "large" $\alpha^{\pm}$. As we will see shortly, when $\alpha^{\pm}$ is large, $(\alpha^{\pm})^2$ becomes a temperature ratio which gives a mea-

sure of the synchronicity of the heat flux; viz., the ratio of the mean velocity squared to the thermal noise. Perfect synchronicity would correspond to $\alpha \to \infty$, and $\langle v_z\rangle^2_{\subset\pm} = \langle v_z^2\rangle_{\subset\pm}$. Values of $(\alpha^\pm)^2$ of $\frac{1}{2}$, 1, and 2 give the area under the Gaussian which is cut off by the ordinate as 16%, 8%, and 3%, respectively, and fix the ratios of $\langle v_z\rangle^2_{\subset\pm}/\langle v_z^2\rangle_{\subset\pm}$ at approximately $\frac{1}{2}$, $\frac{2}{3}$, and $\frac{4}{5}$, respectively. Thus, even for an α of 1, it is a fair approximation to take $\zeta^\pm$ as the area under the entire Gaussian curve for the $\pm$ distribution. Then (51) becomes

$$\zeta^\pm(\vec{\mathbf{r}}) = [\pi^{3/2} m^3/h^3 \lambda_3^r (\lambda_3^\pm)^{1/2}] \exp[-\lambda_1 + (\lambda_2^\pm)^2/4\lambda_3^\pm],$$

$$\zeta = \zeta^+ + \zeta^- \qquad (63)$$

and (28) yields then

$$-\lambda_2^\pm/2\lambda_3^\pm = \langle v_z\rangle_{\subset\pm}(\vec{\mathbf{r}}), \qquad (64)$$

$$1/2\lambda_3^\pm = \langle (v_z - \langle v_z\rangle_{\subset\pm})^2\rangle_{\subset\pm}(\vec{\mathbf{r}}), \qquad (65)$$

$$1/\lambda_3^r = \langle v_r^2\rangle_{\subset\pm}(\vec{\mathbf{r}}). \qquad (66)$$

Temperature

Using the previous definition of experimental temperature, we get

$$(3k/m)\, T^\pm(\vec{\mathbf{r}}) = 1/2\lambda_3^\pm + (\lambda_2^\pm/2\lambda_3^\pm)^2 + 1/\lambda_3^r, \qquad (67)$$

which could be measured by an appropriate one-sided thermometer, say a small thermistor, insulated on one side, and appropriately streamlined. A cylindrical thermometer, on the other hand, will measure a mean local energy per particle, and a mean temperature, which can be written

$$\overline{T}(\vec{\mathbf{r}}) = (\langle n\rangle_+/\langle n\rangle)(T_+ - T_-) + T_-. \qquad (68)$$

The net local density, $\langle n\rangle(\vec{\mathbf{r}})$, could be measured, e.g., by absorption spectrometry, so that (68) affords a computation of the means $\langle n\rangle_\pm(\vec{\mathbf{r}})$. We note that if $\langle n\rangle_+ = \langle n\rangle_-$, then (40) necessitates $\langle v_z\rangle_{\subset+} = -\langle v_z\rangle_{\subset-}$, so that $(3k/m)(T^+ - T^-)$ is then just $1/\lambda_3^+ - 1/\lambda_3^-$, which by (47) is nonzero. Thus we cannot have both $\langle n\rangle_+ = \langle n\rangle_-$ and $T_+ = T_-$.

Pressure

As we saw earlier, the off-diagonal terms for the pressure tensor vanish for this system. The diagonal terms are readily found in terms of the λs by (64)–(66), and related to the experimental temperatures via (67), providing equations of state.

It is of greater interest to consider a "vector pressure"; that is, the resultant *force* at a point due to momentum transport, rather than the usual scalar pressure. We consider the change of momentum within some small volume bounded by the planes z and $z+\Delta z$, over some small, fixed time interval Δt. In our case, this will be due entirely to the thermal flux of momentum across the two planes. The flux *out* of the volume will be given by the usual integrals over the complete half-velocity ranges[27]; i.e., the momentum transported out of the volume is

$$\Delta t\left(\int_0^\infty N m v_z^2 \sigma_1(\vec{\mathbf{v}}, z+\Delta z)\, d\vec{\mathbf{v}}/\hbar^3 - \int_{-\infty}^0 N m v_z^2 \sigma_1(\vec{\mathbf{v}}, z)\, d\vec{\mathbf{v}}/\hbar^3\right) \equiv \Delta t \pi R^2 [P^+{}_{zz}(z+\Delta z) - P^-{}_{zz}(z)]. \qquad (69)$$

To get the flux into the volume, however, we must exclude all those particles crossing the planes with velocities large enough to take them completely through the volume element in the chosen time interval Δt. (We can ignore collisions if we take Δz smaller than the mean free path.) The change in momentum due to the fluxes from the rest of the body of the gas into this element is therefore given by

$$\Delta t\left(\int_0^{\Delta z/\Delta t} N m v_z^2 \sigma_1(\vec{\mathbf{v}}, z)\, d\vec{\mathbf{v}}/\hbar^3 - \int_{-\Delta z/\Delta t}^0 N m v_z^2 \sigma_1(\vec{\mathbf{v}}, z+\Delta z)\, d\vec{\mathbf{v}}/\hbar^3\right) \equiv \Delta t \pi R^2 [P^+{}_{zz}(z) - P^-{}_{zz}(z+\Delta z)]. \qquad (70)$$

To get the momentum change at a *plane* during Δt, we shrink Δz to 0, whence the contribution (70) vanishes. The instantaneous rate of change of momentum is then given solely by (69) for $\Delta z = 0$, divided by Δt, and evaluated in the limit $\Delta t \to 0$; i.e.,

$$\int_0^\infty N m v_z^2 \sigma_1(\vec{\mathbf{v}}, z)\, d\vec{\mathbf{v}}/\hbar^3 - \int_{-\infty}^0 N m v_z^2 \sigma_1(\vec{\mathbf{v}}, z)\, d\vec{\mathbf{v}}/\hbar^3 = \pi R^2 [P^+{}_{zz}(z) - P^-{}_{zz}(z)] \qquad (71)$$

is the force acting across the plane at z due to momentum transport.[29] The pressure *difference* across this plane is then just $1/\pi R^2$ times this. If we were to place a thin vane of some thermally insulating material at z, then if we assume elastic collisions, and that the vane does not seriously perturb the velocity distribution of the impinging particles, the pressure difference across this vane should be exactly twice the value gotten from (71); viz., across vane,

$$P^+{}_{zz}(\vec{\mathbf{r}}) - P^-{}_{zz}(\vec{\mathbf{r}}) = 2(\langle n\rangle_+ \langle v_z^2\rangle_{\subset+} - \langle n\rangle_- \langle v_z^2\rangle_{\subset-}). \qquad (72)$$

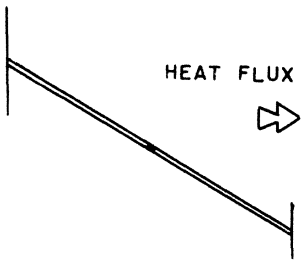

FIG. 3. Torsion pendulum for measuring pressures in the absence of a pressure gradient.

Owing simply to symmetry considerations, all other components of the "pressure vector" across the z^-, or any other, plane must vanish in the interior of the gas. By the requirement of stationarity, we do not have a time-rate of change of *anything*, including momentum, whence (71) and (72) must also vanish, for the present system. This result could be tested for by using a torsion pendulum such as illustrated in Fig. 3, from a top view. This consists of a small rod to which vanes of unequal areas have been affixed at oblique angles. This is then suspended from its center of mass by a fine quartz thread. According to the preceding, such a pendulum should not turn. This, of course, is not a test of the present theory, since the argument just given is quite general and independent of the specific form of the probability density. However, we can derive from this a condition which the information theoretic analysis specifically requires. If we had $\langle n\rangle_+ = \langle n\rangle_-$, (40) and the vanishing of (72) would require, as discussed in the section on temperature, that $\lambda_3^+ = \lambda_3^-$, contrary to (47). Since $T_+ = T_-$ would necessitate that $\langle n\rangle_+ = \langle n\rangle_-$, we can, then, rule out these possibilities entirely; i.e., we must have *both*

$$\langle n\rangle_+ \neq \langle n\rangle_- \quad \text{and} \quad T_+ \neq T_- \tag{73}$$

at all planes z.

Heat flux

In the absence of $\lambda_4^\pm$, we can write (43) in the form $\phi_E = \phi_E^+ + \phi_E^-$, with $\phi_E^\pm = (mN/2\xi)\partial^2\zeta^\pm/\partial\lambda_2^\pm\partial\lambda_3^\pm$; substitution of (63)–(65) then gives

$$\phi_E^\pm = \tfrac{1}{2}m\langle n\rangle_\pm\langle v_z\rangle_{\subset\pm}(3\langle v_z^2\rangle_{\subset\pm} - 2\langle v_z\rangle_{\subset\pm}^2). \tag{74}$$

It does not seem possible to draw many conclusions on the ± populations solely on the basis of (74). It is possible to get sizeable energy fluxes, say corresponding to He at approximately room temperature, in a gradient of 10° K/cm, and still have the $T^\pm$ differ by as little as 10^{-4} °K. In this case however, the bimodality of Fig. 2 would be very pronounced, the peak separation being roughly 10^5 cm/sec. Many other combinations are possible.

Entropy

By using (63)–(66), (7), (29), and (30) in (33) and (34), we get for the local entropy $\mathfrak{s}(\vec{r}) = \mathfrak{s}^+(\vec{r}) + \mathfrak{s}^-(\vec{r})$, where

$$\mathfrak{s}^\pm(\vec{r}) = k\{\langle n\rangle_\pm \ln[(m^2\pi e/h^2)^{3/2}(1/\lambda_3^r)(1/\lambda_3^\pm)^{1/2}] - \langle n\rangle_\pm \ln(\langle n\rangle_\pm/e)\} \tag{75}$$

or

$$\mathfrak{s}^\pm(\vec{r}) = k\mathfrak{n}^\pm(\vec{r})\ln\{4e^{5/2}(m\pi/h^2)^{3/2}(\mathfrak{n}^\pm(\vec{r}))^{-5/2}\mathfrak{u}_r^\pm(\vec{r}) \times[\mathfrak{u}_z^\pm(\vec{r}) - m(\mathfrak{v}_z^\pm(\vec{r}))^2/2\mathfrak{n}^\pm(\vec{r})]^{1/2}\} \tag{76}$$

and we have defined the extensive variables $\mathfrak{n}^\pm \equiv \langle n\rangle_\pm$, $\mathfrak{u}_r^\pm \equiv \frac{1}{2}m\mathfrak{n}^\pm\langle v_r^2\rangle$, $\mathfrak{u}_z^\pm \equiv \frac{1}{2}m\mathfrak{n}^\pm\langle v_z^2\rangle_{\subset\pm}$, and $\mathfrak{v}_z^\pm \equiv \mathfrak{n}^\pm\langle v_z\rangle_{\subset\pm}$.

By a judicious application of Gibb's theorem for the discrete and continuous cases to the various terms in the integral of (76) over the volume, one can see that the spatial inhomogeneity and partitioning act to reduce the total entropy of the system. More rigorously, and simply, if we compare against a system with the same total energy and particle number—e.g., if we drop adiabatic shields around the system after it has attained the stationary state—then the entropy of the equilibrium system that results will be larger, by (17). Thus the energy flux acts to organize the system.[30] It is perhaps worth commenting that this does *not* follow from the second law of thermostatics. This law, when stated in terms of the entropy, characterizes the direction of evolution between equilibrium states, upon the removal of a constraint.[31] We are presently concerned with the direction of evolution of a highly nonequilibrium system which is allowed to go to equilibrium. Intuition aside, nothing meaningful can be said about this until we have a perfectly general and meaningful definition of entropy, which is exactly what information theory gives us. Since, for the present system, all the information (except particle number, which is constant) is in the one-particle distribution function, the H theorem is applicable here, and also characterizes the direction of evolution upon isolation.[27] What would be much more interesting, of course, would be the direction of evolution between different nonequilibrium, nonisolated, states upon the manipulation of various external constraints. The H theorem can tell us nothing about this. Methods such as illustrated here in (16) and (17) at least allow us

a means of comparison once we know what the two states in question are, and are perfectly general.

The entropy minimum for this system cannot be found by a variational procedure, since it occurs at the end of the interval of definition of δ, for $\delta(\vec{r})=0$. By means of (64), (65), and (76), this can be translated into a limiting relation on the two measures of "synchronicity," $(\alpha^{\pm})^2$, defined above.

We close this section by observing that the limiting case just treated may actually never occur. Nothing thus far found indicates that this is necessarily the high-flux limit. We have chosen it merely for ease of treatment.

VI. COMPARISON WITH WORK OF OTHERS

Holway[22] has dealt specifically with the application of information theory to kinetic theory and the evaluation of the collision term in the Boltzmann equation. He went so far as to introduce second moments, corresponding to the components of our (21) for $l=2$, which were not collisional invariants, as his informational constraints. But he did not consider more general moments, or fluxes, and his selection appears to have the same sort of arbitrariness as the more conventional approaches. He also appears to regard the resulting probability densities which he obtains less seriously than we intend here. He does not attach any physical significance, e.g., to the corresponding entropies.

On the *non*-information-theoretic end of things, there has already been a fair amount of work using half-range probability densities to treat this system. This splitting was first introduced at the boundaries,[2-4] and gradually worked its way into the interior of the gas.[5-10] It was early recognized that the probability density must be nonanalytic in v_z in order to satisfy the boundary conditions, and this limiting behavior at the boundaries seems to be often taken as outright discontinuity.[6-8,13] While we have assumed throughout the present paper that σ must at least be continuous, we acknowledge that this is a question to be considered. If we take the point of view that σ_1 is proportional to a particle density, it would seem that *this* must be continuous. Whether that continuity is best represented *in* σ_1, or by drawing some very steep line between $\sigma_1(dv_z)$ and $\sigma_1(-dv_z)$, for small dv_z, is open to debate.

It seems that Gross *et al.* were the first to propose, as a computational device, the use of "half-range distribution functions" throughout the body of a gas for a certain class of problems, and gave a thorough treatment of this general method.[5] This was then applied to the heat-flow problem considered here by several workers.[6-13] Gross and Ziering[6] and Ziering[7] use a linearized form of σ_1 corresponding to the expansion introduced here, for small thermal gradients, in the second part of Sec. V. Translated to our notation, they have treated the two cases where the nonvanishing λs are the sets $\{\lambda_1^+\neq\lambda_1^-,\ \lambda_3^+=\lambda_3^{r+}\neq\lambda_3^-=\lambda_3^{r-}\}$ and $\{\lambda_1^+\neq\lambda_1^-,\ \lambda_2^+\neq\lambda_2^-,\ \lambda_3^+=\lambda_3^{r+}\neq\lambda_3^-=\lambda_3^{r-},\ \lambda_4^+\neq\lambda_4^-\}$, and solved the Boltzmann equation to obtain, among other things, $\langle n\rangle(z)$, $\overline{T}(z)$, and $\phi_E(l/L,\Delta T)$. Their second moment treatment, and the treatments by several others,[10-13] do not seem to be in accord with the information on scattering by surfaces,[16-20] which would require $\lambda_2^{\pm}\neq 0$, at least at the walls.

Lees[9] and Lees and Liu[10] have treated the case for small ΔT, for both planar and cylindrical geometry, and introduce a σ_1 which appears to be the same as our second moment approximation,[28] illustrated here as Fig. 2, although the parameters they choose are quite different than ours. In actual calculations, however, they revert to a simpler form, corresponding to $\lambda_2^{\pm}=0$.

Of the existing treatments of this problem, these "half-range" distributions have provided the closest agreement with the experiments that have been performed for small ΔTs and Knudsen numbers ranging from the free molecule limit to continuum[14,15]; e.g., the second-moment distribution of Gross and Ziering, and that of Lees, give the density distribution to within 3% of the experimental curve, and the heat flux to within 2% of the experimental values.[15] Unfortunately, there seems to be a dearth of data for higher temperature gradients, although theoretical results exist.[9,11,12] From our point of view, the improved agreement with experiment, over more involved calculations, such as Wang Chang and Uhlenbeck's[3] (which is still quite good) can be expected to carry over to arbitrary ΔT, and has its reason in the distribution functions being closer to the truth. However, these workers seem reticent to ascribe to these σ's full status as velocity distributions, and emphasize rather the computational advantage. Lees and Liu,[10] e.g., state that "... it must be stressed that each individual function $[\sigma_1(\vec{v}|\pm)]$ has no explicit physical significance in general."

Krook has written a particularly lucid account[8] of the half-range treatment of the heat-flow problem, and goes so far as to identify moments of the form $\langle V^l\rangle_{\pm}$ as the "state variables" of the system, rather than merely intermediates in the calculation of net observables. He gives a nice discussion of the effect of collisions on the two-sided nature of the distribution function, offering some argument for what seems to be a widely-held assumption,[3,4,7,12] that collisions will remove this

two sidedness, at least in the density-independent flux range (Knudsen number, $l/L \ll 1$). Although we would agree that the effect of collisions is in this direction, we cannot agree that for a system of a given density, collisions will "attenuate" the differences that exist at the walls. Even for small ΔTs, the near-equilibrium calculation given here, and those by Gross and Ziering[6] of the λs in terms of collision integrals for hard spheres, show that the difference between these $\lambda^{\pm}s$ is *quantitatively preserved* at all distances between the walls. From these same calculations, we furthermore see that, even for $l/L \sim 10^{-4}$, such differences must remain finite, since the heat flux is directly proportional to them.

Some of the major differences between the conventional approaches to this problem and the development given here are summarized in Table I.

VII. CONCLUSION AND EXHORTATION

The present paper has had the limited objective of bringing the statistical-mechanical analysis to the point where the phenomenology of this nonequilibrium system could begin to be developed. Our principal results are (44)–(49), (51), and for those who would like to see what a specific far-from-equilibrium entropy looks like—just one—we have Eq. (76). What is perhaps the most distinguishing feature to emerge from the analysis, and which we might expect to carry over to other nonequilibrium systems, is the development of structure within the system, so that the new state variables are no longer independent, as in equilibrium. This is exactly analogous to the effect of inhomogeneity: In that case, a single set of global variables no longer suffice to adequately define the thermal state of the system, but rather each such variable becomes a prescribed function of position. Just so here, the state variables, besides being functions of position, depend further on the region of the microstate space being considered. In retrospect, it seems rather curious that in treating nonequilibrium systems people would be so persistent in *precluding* the possibility of such variety and functional organization within the system.

The critical reader is liable to have two, related, complaints about the preceding analysis: It is neither quantitative, nor exhibits any specific account of the dependence on density. We have said, for example, that the $\lambda^{\pm}$'s must be different, but not *how* different. Again, aside from the perturbation calculation made here, where we introduced an *ad hoc* argument about the dependence of $T^{+}(z) - T^{-}(z)$ on the mean free path, specific mention of the effects of density are absent. In that specific case, the argument allows us to show that for the free molecule limit, where $l = L$, the distance between the walls, our analysis gives $\phi_E \propto \langle n \rangle$, as it should, while in the high

TABLE I. Comparison of approaches to the heat-flow problem.

Conventional approaches	Present information-theoretic approach
Start with Boltzmann's equation, whose rigorous derivation is from Liouville's theorem. Strictly valid only for isolated systems.	Completely independent of Liouville's theorem, or Boltzmann's equation.
Approximate collision term in Boltzmann equation by linear operator. Good only for near equilibrium.	No approximations used for ideal-gas model. More general applicability is limited only by effect of collisions on the informational constraints; i.e., the potential contribution to the fluxes.
Standard treatments generally involve one or more additional assumptions of near equilibrium along the way.	No limitations as to the degree of nonequilibrium.
Chapman-Enskog procedure good only for small ΔTs, continuous density, and $\langle n \rangle$, $\langle \vec{\mathbf{v}} \rangle$, and $\langle u \rangle$, as state variables.	Good for arbitrary ΔT, and arbitrary density, so long as the kinetic contribution to the fluxes dominates. σ_1 is found in terms of the new, nonequilibrium state variables.
Grad's method assumes σ_1 a local Maxwellian, and expands about this equilibrium form. Truncation of the expansion is arbitrary.	σ_1 is limited only by what is known. It is found exactly, in simple finite form.
When half-space probability densities are used, they are introduced *ad hoc*, as a computational device; not interpreted as a particle density.	$\sigma_1(\vec{\mathbf{v}}\,|\pm)$ shown to be the necessary and natural decompositon of σ_1, relating it to the new thermal state variables of the system. These $\sigma_1(\vec{\mathbf{v}}\,|\pm)$ give the particle densities.

density limit that $l \propto \langle n \rangle^{-1}$, we have ϕ_E independent of $\langle n \rangle$, as it is known to be. The model chosen here, a gas of "hard points," does not have an excluded volume, and one might well expect that, aside from some effect of the randomization of the components of momentum upon collision, density does not make any difference, and the gas will always be essentially a Knudsen gas, in which case all $\vec{r}$ dependence of the observables vanishes. But the model chosen was for convenience, so that the expressions (39)–(43), with σ given by (24) and (44), would be exact. The limitation on applying the present results to any particular real gas rests entirely on how well the means given by (39)–(43) are indeed approximated by kinetic terms. For the inert gases, this should be a pretty good approximation, and for He, which behaves very much like a hard-sphere gas over wide range of temperature, it should be quite good, up to as much as 10 atm. The further refinement of the theory to give the quantitative dependence of the observables on density and position will of course necessitate an analysis of at least hard-sphere collisions. This, and the subsequent development of the phenomenology of this system, will be the subject of a later paper.

In conclusion, we make some general remarks. The development of the body of physics labeled nonequilibrium thermodynamics has to the present time been dominated by an understandable leaning on the much easier, and better understood, equilibrium theory, and an attempt to build from there, for example, by perturbation methods, and assumptions of *local* equilibrium. If one studies real nonequilibrium systems—any chemical reaction, or living systems, as the supreme example—it does not take long to be convinced that the difference is deeper and more thoroughgoing than these methods will ever be able to account for.

In his 1957 papers,[23] Jaynes clearly spelled out the possibility of developing a completely general statistical mechanics from the information theoretic standpoint. Since that time, it seems that all work has been either abstract and formal, on the one hand, or the treatment of specific systems which can already be handled by other means, on the other. Largely, the problem seems to be that even those utilizing the information theoretic approach are unwilling to take it seriously in its generality, and instead buttress their derivations with Liouville's theorem at every opportunity. If there is any question as to the applicability of Liouville's theorem to real, "isolated" systems,[32] there should be no question as to its *inapplicability* to real systems whose *raison d'etre* is that they are *not* isolated. The more practical minded, on the other hand, seem suspicious of the physical meaning of the information theoretic entropy, so that this has been relegated to the exclusive study of mathematicians.[33]

In this, and the previous[28] paper we have attempted to bridge this gap in the simplest instance thought possible: a one-dimensional heat flow through a stationary ideal gas. The phenomenology as thus far developed already yields a decisive contradiction with the conventional assumptions for this, and like, systems, in (60) and (61), and again in (73); according to the present analysis, the "kinetic temperature"[27] is *different* for molecules moving in the $\pm z$ directions, respectively, in these two cases. This can be readily checked experimentally, as discussed above. Again, at the microscopic level, the velocity distribution, examples of which are shown in Figs. 1 and 2, can be determined by several experimental methods,[21] as was previously suggested.[28] While the significance of the specific results found here hinges entirely on such experimental verification, the analysis should be of heuristic value, in any event, for those who think that information theory is the correct foundation for statistical mechanics.

The major need at present, for the development of *any* kind of truly nonequilibrium thermal physics, appears to be for experimental work to determine the new thermodynamic variables, and relations between them, for specific, well-defined nonequilibrium systems.

ACKNOWLEDGMENTS

The author would like to thank Dr. John L. Oncley and Dr. Harold J. Morowitz for their valuable interest, encouragement, and criticism of the above investigation.

---

*Research supported in part under National Institute of Health Training Grant No. GM1355.

†National Institute of Health predoctoral trainee.

[1]In particular, the following conventions will be used throughout the paper. { } will be used exclusively to denote a set of quantities. Expressions like $f[x]$ indicate a functional of x. Expressions like $\langle f \rangle(\vec{r})$ denote a local mean. For brevity, we will often drop specific reference to the $\vec{r}$ dependence from this, and other, expressions; the great majority of the discussion is in terms of local quantities. On the few occasions we will have to speak of global means, they will be indi-

cated by an asterisk, as in $\langle f \rangle^*$, so there should be no ambiguity. In integrals, quantities like $d\vec{r}$ denote differential volume elements. By differentiation with respect to a vector, $df/d\vec{A}$, is meant the vector $(df/dA_x, df/dA_y, df/dA_z)$. Other notation is defined in the text.

[2]C. S. Wang Chang and G. E. Uhlenbeck, Transport Phenomena in Very Dilute Gases, University of Michigan Engineering Research Institute Report No. CM-579, UMH-3-F, 1949 (unpublished).

[3]C. S. Wang Chang and G. E. Uhlenbeck, The Heat Transport Between Parallel Plates as Functions of the Knudsen Number, University of Michigan Engineering Research Institute Project M999, 1953 (unpublished).

[4]H. M. Mott-Smith, New Approach in Kinetic Theory, Lincoln Laboratory Report No. V2, 1954 (unpublished).

[5]E. P. Gross *et al.*, Ann. Phys. (N.Y.) 1, 141 (1957).

[6]E. P. Gross and S. Ziering, Phys. Fluids 2, 701 (1959).

[7]S. Ziering, Phys. Fluids 3, 503 (1960).

[8]Max Krook, J. Fluid Mech. 6, 523 (1959).

[9]L. Lees, A Kinetic Theory Description of Rarefied Gas Flows, GALCIT Hypersonic Research Project Memo. No. 51, 1959 (unpublished).

[10]L. Lees and Chung-Yen Liu, Phys. Fluids 5, 1137 (1962).

[11]Y. S. Lou and T. K. Shih, Phys. Fluids 15, 785 (1972).

[12]M. L. Lavin and J. K. Haviland, Phys. Fluids 5, 274 (1962).

[13]K. Frankowski *et al.*, Phys. Fluids 8, 245 (1965).

[14]A. Dybbs and G. S. Springer, Phys. Fluids 8, 1946 (1965).

[15]W. P. Teagan and G. S. Springer, Phys. Fluids 11, 497 (1967).

[16]H. Jakus and F. C. Hurlbut, in *Rarefied Gas Dynamics*, edited by L. Trilling and H. Y. Wachman (Academic, New York, 1969), Suppl. 5, Vol. II, p. 1171. There are many other works of relevance, too numerous to list, in the supplements of these symposia; we note especially Suppl. 3, Vol. II.

[17]Terrell Marshall Jr., Ref. 16, p. 1079.

[18]R. M. Logan and R. E. Stickney, J. Chem. Phys. 44, 195 (1966).

[19]F. O. Goodman, J. Phys. Chem. Solids 26, 85 (1965).

[20]S. Nocilla and M. G. Chiado-Piat, Ref. 16, p. 1069.

[21]See, e.g., E. P. Muntz, in *Rarefied Gas Dynamics*, edited by J. H. de Leeuw (Academic, New York, 1966), Suppl. 3, Vol. II; also, *Rarefied Gas Dynamics*, edited by C. L. Brundin (Academic, New York, 1967), Suppl. 3, Vol. II.

[22]Lowell H. Holway, Jr., Phys. Fluids 9, 1658 (1966).

[23]E. T. Jaynes, Phys. Rev. 106, 620 (1957); Phys. Rev. 108, 171 (1957).

[24]The formulation in terms of functionals is given by E. T. Jaynes, in *Statistical Physics III*, edited by K. W. Ford (Benjamin, New York, 1963), and applied by B. Robertson, Phys. Rev. 144, 151 (1966) and Phys. Rev. 160, 175 (1967).

[25]See, e.g., I. M. Gelfand and S. V. Fomin, *Calculus of Variations* (Prentice-Hall, Englewood Cliffs, N. J., 1963).

[26]Conditional probabilities, conditional entropies, and much other interesting material is treated in S. Watanabe, *Knowing and Guessing* (Wiley, New York, 1969).

[26a]*Note added in proof.* The vanishing of $\langle A_j \rangle$ is meant here in the strong sense that this observable does not exist for any manipulation of the external constraints acting on the system (see Sec. IV). It should be clear that it is the *constancy* of $\langle A_j \rangle$ with respect to all the intensive variables, not the zero value of that constant, that leads to (19).

[27]See, e.g., J. O. Hirschfelder *et al.*, *Molecular Theory of Gases and Liquids* (Wiley, New York, 1954) and S. Chapman and T. G. Cowling, *The Mathematical Theory of Non-Uniform Gases* (Cambridge U.P., Cambridge, England, 1958).

[28]A. B. Corbet and H. J. Morowitz, Phys. Rev. A 6, 2298 (1972). Each λ of the present paper is minus times the corresponding λ of this previous paper. We have made this change in order to be in accord with the convention adopted by most of the writers on this subject, including Jaynes.

[29]Clearly the order in which we take limits here is important. If we first divide by Δt and let $\Delta t \to 0$, the momentum fluxes at each plane go to the usual full-range integrals, given by (41). We then have $P_{zz}(z) - P_{zz}(z+\Delta z) \to 0$, as $\Delta z \to 0$. On *a priori* grounds we see that this cannot be true as a *general* result. On the other hand, giving the time-rate of change of momentum at a plane as $P_{zz}(z)$, as is sometimes done, can also be ruled out on *a priori* grounds since this quantity is intrinsically positive, and is not even a vector. It would not be appropriate to our present objectives to give a more general or rigorous derivation at this place. The reader is liable to grant (72) even if he disagrees with (71).

[30]For a much more general discussion of this phenomenon, see H. J. Morowitz, *Energy Flow in Biology* (Academic, New York, 1968).

[31]See, e.g., H. B. Callen, *Thermodynamics* (Wiley, New York, 1960).

[32]See, e.g., H. S. Robertson and M. A. Huerta, Am. J. Phys. 38, 5 (1970).

[33]See, e.g., W. Parry, *Entropy and Generators in Ergodic Theory* (Benjamin, New York, 1969).