

Generalized Statistical Mechanics

PETER G. BERGMANN

Department of Physics, Syracuse University, Syracuse 10, New York

(Received July 26, 1951)

In this paper, a start is made toward the development of a statistical mechanics that will be suited to the treatment of dynamic changes in thermodynamic systems and which, at the same time, will be in the appropriate form for a relativistically invariant theory. The formal device consists of making the time coordinate one of the canonical variables. As a result, stationary ensembles no longer occupy a privileged position among all conceivable Gibbs ensembles, and it becomes necessary to redefine and to reformulate most statistical and thermodynamic concepts. This has been done with the concept of a canonical ensemble, with entropy, temperature, heat flux, and performance of work. With a suitable definition of the entropy of an individual system, a new formulation of the H -theorem is provided, which in turn leads to a formulation of the Second Law.

I. INTRODUCTION

EARLIER attempts to formulate a consistent relativistic thermodynamics¹⁻⁵ have usually been concerned with giving the concepts of entropy, temperature, and heat a relativistically invariant significance. Except for the very early work by Einstein and Planck, these attempts have been largely on the so-called phenomenological level, without any reference to the underlying statistical concepts. As a result, many of these papers appear rather formal in approach, achieving their results primarily by the method of seeking Lorentz-invariant modifications of the concepts familiar in ordinary thermodynamics. Eckart⁵ has pointed out that ordinary thermodynamics is fragmentary in so far as it deals only with equilibrium situations and with quasistatic processes. Essentially dynamical processes are outside its scope. On the other hand, relativistic theories become necessary in physics only if velocities are involved which by ordinary laboratory standards are very large. While it is always possible to introduce a local framework in which any one velocity is reduced to zero, a truly relativistic theory is likely to give novel and therefore interesting answers only if velocities are involved that are very large relative to each other. But if two streams of matter interact with each other whose relative velocity is comparable to that of light, then no observer can describe that interaction as a quasi-static process, and the approach of formal invariantization breaks down.

These considerations lead inescapably to the conclusion that the development of relativistic thermodynamics can be tackled successfully only as part of a more comprehensive program, to develop a theory of heat that is independent of the concept of thermal equilibrium. Such a theory we shall call a generalized theory of heat, for want of a better expression. Such a

development is bound to lead to a much more complex theory than the conventional thermodynamics, for the simple reason that a stationary situation has fewer degrees of freedom than one in which time derivatives of the thermal variables also enter into the description. The conventional theory is quite trustworthy in its assertion that a violent development leading from one stationary situation to another stationary situation is bound to increase the entropy. But if there is neither a stationary initial situation nor a stationary terminal situation, then we cannot assign entropies to two states succeeding each other in the course of time without going beyond the confines of conventional thermodynamics. It is not that conventional thermodynamics gives the wrong answers; it fails to give any answer at all.

If we are to generalize the conventional theory of heat, then we must, in this writer's opinion, start out with the fundamentals, and that means not with thermodynamics, but with statistical mechanics. Hence this paper is devoted to the development of a statistical mechanics independent of the notion of the stationary ensemble. Formally, we shall proceed by depriving the time coordinate of its privileged role vis-a-vis the usual canonical coordinates. While the formal method is well known and amounts to the addition of two further dimensions to the ordinary phase space (time and its canonical conjugate, the negative Hamiltonian), the description of an ensemble in this new phase space in terms of the density of its member systems is complicated by the fact that all the systems are constrained to move on one hypersurface, and furthermore, each individual system is no longer represented by a point but by a curve, its trajectory. Thus, the ensemble must be described in terms of the density of a field of curves on a hypersurface. Fortunately, such a description in terms of geometric invariants is quite feasible and leads to the introduction of a single invariant quantity which for all practical purposes assumes the role of the usual density in phase space.

The development of this description occupies most of the following two sections. The invariant description

¹ A. Einstein, *Jahrbuch d. Radiaktivität und Elektronik* 4, 441 (1907).

² M. Planck, *Preuss. Akad. Wiss. Berlin, Sitz. Ber.* 29, 542 (1907).

³ M. Planck, *Ann. Physik* 26, 1 (1908).

⁴ R. C. Tolman, *Relativity, Thermodynamics, and Cosmology* (Oxford University Press, London, 1934).

⁵ C. Eckart, *Phys. Rev.* 58, 919 (1940).

of an ensemble of quantum-theoretical systems is also indicated. From then on, we shall reformulate the concept of the canonical ensemble, and with it that of the ensemble entropy. We shall find that in general a single temperature is not sufficient to characterize a canonical ensemble, nor can the interaction of a thermodynamic system with its surroundings be adequately described in terms of a scalar "heat transfer" and another scalar "performance of work." But there are natural generalizations of these all-important quantities, and there exists a relationship between the generalized heat transfer, multiplied by the generalized reciprocal temperature, and the increase in entropy.

In this generalized statistical mechanics, the development of a relativistic, i.e., Lorentz-invariant, theory depends solely on the availability of a relativistic Hamiltonian. Such Hamiltonians are, of course, known and of particular importance in covariant field theories (such as the theory of the electromagnetic field). Thus, relativistic thermodynamics requires no additional concepts.

II. PARAMETRIZATION

In the ordinary formulation of analytical mechanics, the time t is not a dynamic variable while the generalized coordinates q_k are. In Lorentz transformations, on the other hand, the time appears symmetrically with the spatial coordinates, and the (negative) energy with the linear momentum components. It is, therefore, desirable, to modify the usual analytical mechanics in such a manner that the privileged position of the time coordinate and its canonical conjugate, the (negative) energy, is destroyed. Similar comments apply to the analogous quantum-mechanical situation.

A. Classical Mechanics

Consider a system with f degrees of freedom whose generalized coordinates shall be designated by q_k and their canonical conjugates by p_k , k taking values from 1 to f . Let the Hamiltonian be a known function $H(q_k, p_k, t)$. It is not necessary to assume for what follows that the system be conservative, i.e., that the Hamiltonian be explicitly independent of t . In the phase space S with the coordinates q_k, p_k , the trajectory of the system is then a single curve which depends in a definite manner on the time t . In other words, if we solve the equations of motion, the resulting functions $q_k(t), p_k(t)$ represent not merely a convenient parametric representation of the mechanical trajectory, but t is a physically significant variable of its own, though not a coordinate of our present phase space.

It is well known that the canonical equations of motion are the Euler-Lagrange equations of a variational principle in phase space,

$$\delta S = 0, \quad S = \int [\sum_k p_k dq_k/dt - H(q_k, p_k, t)] dt. \quad (2.1)$$

If we wish to "symmetrize" with respect to the variable t , we may interpret the variational principle as one in a $(2f+2)$ dimensional space, possessing the coordinates q_k, p_k, t, \bar{E} (the latter being equal to $-H$). The variation in that higher-dimensional space, which we shall call the "expanded phase space," is, however, restricted by a subsidiary condition,

$$\bar{H} \equiv H(q_k, p_k, t) + \bar{E} = 0. \quad (2.2)$$

The variational principle itself can be formulated conveniently if we introduce an arbitrary parameter $\theta(q_k, p_k, t, \bar{E})$, subject only to the inequality

$$\partial\theta/\partial t > 0, \quad (2.3)$$

and takes the form

$$\begin{aligned} \delta S = 0, \quad S &= \int \left(\sum_{k=1}^f p_k \frac{dq_k}{d\theta} + \bar{E} \frac{dt}{d\theta} \right) d\theta \\ &= \int \left(\sum_{\kappa=1}^{f+1} p_\kappa \frac{dq_\kappa}{d\theta} \right) d\theta = \int \left(\sum_{\kappa=1}^{f+1} p_\kappa dq_\kappa \right), \end{aligned} \quad (2.4)$$

where we have used the abbreviating notation

$$q_{f+1} = t, \quad p_{f+1} = \bar{E}; \quad \kappa = 1, \dots, f+1. \quad (2.5)$$

The variational principle (2.4) with the subsidiary condition (2.2) in the expanded phase space is equivalent to the principle (2.1) in the original phase space. The canonical equations of motion are obtained by the method of Lagrange's multipliers and come out as

$$dq_\kappa/d\theta - \alpha \partial \bar{H} / \partial p_\kappa = 0, \quad dp_\kappa/d\theta + \alpha \partial \bar{H} / \partial q_\kappa = 0. \quad (2.6)$$

The multiplier α cannot be determined uniquely. Its arbitrariness reflects the fact that the parameter θ is arbitrary, and that this new formalism is invariant with respect to arbitrary parameter transformations. Instead of keeping the multiplier in the equations of motion (2.6), we shall incorporate it in the subsidiary condition and set, from now on.

$$\begin{aligned} dq_\kappa/d\theta = \partial H / \partial p_\kappa, \quad dp_\kappa/d\theta = -\partial H / \partial q_\kappa, \\ H = 0, \end{aligned} \quad (2.7)$$

where H is an arbitrary function of \bar{H} and $(2f+1)$ canonical coordinates, so that $H=0$ where $\bar{H}=0$, and that for $\bar{H}=0$ the derivative $(\partial H / \partial \bar{H})$ does not vanish.

The canonical transformations in the expanded phase space are all those which will not change the form of the variational integral (2.4). The form of the subsidiary condition is usually changed by the substitution of the new variables for the old ones. In the special case of an infinitesimal canonical transformation, the changes in the coordinate values at any fixed point in the expanded phase space are

$$\delta q_\kappa = \partial \Gamma / \partial p_\kappa, \quad \delta p_\kappa = -\partial \Gamma / \partial q_\kappa, \quad \Gamma = \Gamma(q_\kappa, p_\kappa). \quad (2.8)$$

δH vanishes by definition. But if we consider H as a function of its arguments, then its functional de-

pendence changes, and the left-hand side of the subsidiary condition undergoes the infinitesimal change

$$\delta'H = (\Gamma, H). \quad (2.9)$$

We find that the whole formalism is invariant with respect to infinitesimal transformations which are generated by constants of the motion, i.e., dynamical variables (which may depend on t explicitly) whose value does not change along a trajectory. Incidentally, the Poisson bracket symbol (Γ, H) is to be understood in terms of the expanded phase space,

$$(A, B) \equiv \sum_{\kappa=1}^{f+1} \left(\frac{\partial A}{\partial q_{\kappa}} \frac{\partial B}{\partial p_{\kappa}} - \frac{\partial A}{\partial p_{\kappa}} \frac{\partial B}{\partial q_{\kappa}} \right). \quad (2.10)$$

In Sec. III, we shall construct ensembles in the expanded phase space. Let it be noted that only trajectories satisfying the subsidiary condition may be incorporated in any ensemble that is to be meaningful. On the other hand, any trajectory satisfying the subsidiary condition at one point satisfies the condition everywhere. The subsidiary condition is itself a constant of the motion.

B. Quantum Mechanics

In quantum mechanics, one might think first of parametrizing by starting out from the wave-mechanical picture, introducing the time t as an additional dimension of configuration space. While this procedure is perfectly feasible and leads to the desired results, it suffers from the aesthetic disadvantage that the results obtained are not obviously independent of the representation chosen. That is why we shall adopt a different approach. Without changing the usual definitions of norm, Hermitian operator, etc., we shall introduce, in addition to the usual operators in Hilbert space, the operations of multiplication by, and differentiation with respect to, time, applied in the Schrödinger representation. Considered at first merely as linear operators in the linear vector space of time-dependent Hilbert vectors (i.e., Hilbert vectors as functions of t), these new operators obviously commute with the operators p_k, q_k , and with each other they satisfy the commutation relation

$$[\bar{E}, t] = \hbar/i, \quad \bar{E} \stackrel{S}{=} (\hbar/i) \partial/\partial t. \quad (2.11)$$

It is not possible to extend the definition of the usual Hilbert space so that these new operations become Hilbert operators. That is because the operation $(H + \bar{E})$ also is conjugate to t (satisfies a relation like (2.11)), and therefore as a Hilbert operator cannot have any normalized eigenfunctions. On the other hand, every solution of the Schrödinger equation is a null vector of the operator $(H + \bar{E})$. Hence, \bar{E} and t cannot both become Hilbert operators.

On the other hand, it is quite possible to construct new operations from the set q_k, p_k, t, \bar{E} which as linear

operations satisfy the standard commutation relations. It is appropriate to call the transition to a new set of $(2f+2)$ such operations an expanded canonical transformation. The question arises how among all these linear operations one can characterize the Hilbert operators, which alone can have the mathematical properties that enable them to represent measurable physical quantities.

To this end, we shall consider, among all the conceivable time-dependent Hilbert vectors, the subset of those satisfying the condition

$$(H + \bar{E})\psi = 0, \quad (2.12)$$

i.e., the Schrödinger equation. This equation has a meaning independent of the Schrödinger representation, because the operations t and \bar{E} , while defined originally only in the Schrödinger representation, can be generalized by the proviso that in any other representation they are to be realized by the corresponding transformed operations. We can bring the subset of time-dependent Hilbert vectors consistent with Eq. (2.12) into one-to-one correspondence with the set of ordinary Hilbert vectors, by identifying each ordinary Hilbert vector with the one in the time-dependent subset which is identical with it at the fixed time t_0 . Moreover, on the strength of this identification the concepts of norm and of scalar bracket can be transferred to the subset of time-dependent Hilbert vectors in a unique manner which is independent of the choice of representation. Briefly, the subset forms again a Hilbert space, with all its ordinary properties. An "operation" will be an "operator" in that Hilbert space if it maps members of the subset into the subset again. The operation t , for instance, clearly does not possess this property. It is not an operator in the new Hilbert space.

The new Hilbert space consists of solutions of the Schrödinger equation. An operator in our new sense must be an operation that applied to a solution of the Schrödinger equation produces another solution of the Schrödinger equation. Operations having this property are ordinarily called constants of the motion. In a Heisenberg representation, they are the operators that are constant in time.

The characteristic of a constant of the motion is that it commutes with the operation $(H + \bar{E})$, which henceforth we shall designate by \bar{H} . Since \bar{H} certainly commutes with itself, it is an operator, albeit a trivial one: all quantum-mechanical states are eigenstates belonging to the eigenvalue 0. Ordinary operators, such as "the value of q_1 at the time t_1 " can be converted into constants of the motion by a simple expedient: at all other times we identify with the operator $q_1(t_1)$ that operator $Q(t)$ which has the same bracket (Dirac bracket) value with any two states at the time t that $q_1(t_1)$ has at the time t_1 . The operation t cannot be so extended except in a trivial manner. "The time t at the time t_1 " is simply the c -number t_1 .

These new characterizations of the ordinary tools of

quantum mechanics have the advantage that only invariant formations occur and that they are invariant with respect to the expanded canonical transformations in which the time t is handled symmetrically with other canonical variables. Even without the explicit introduction of the parameter θ , which we used in the classical part of this section, it is clear that t no longer plays a special role.

III. ENSEMBLE DENSITY

In the ordinary construction of a Gibbs ensemble in classical statistical mechanics, the density of the representative points in phase space is described by a variable μ , a function of the canonical variables q_k , p_k and the time t , which satisfies Liouville's theorem

$$\partial\mu/\partial t + (\mu, H) = 0, \quad \int \mu dX = 1. \quad (3.1)$$

The ensemble average of an observable A is given by the integral $\int \mu A dX$. In the quantum-mechanical formulation a similar role is played by a Hermitian operator μ , constructed in such a manner that the ensemble-average expectation value of any observable A is given by the trace $\text{tr}\{\mu A\}$, with $\text{tr}\{\mu\} = 1$.

In the expanded phase space, the individual systems are described not by representative points, but by representative trajectories. Moreover, these trajectories cannot form a field of curves throughout the expanded phase space, but only on the hypersurface given by Eq. (2.2). Our problem is then to introduce in some fashion a quantity describing the density of trajectories on this hypersurface. We shall find again that this density will be a scalar with respect to (expanded) canonical transformations.

To describe the situation, we shall, in addition to the $(2f+2)$ canonical coordinates q_k , p_k , introduce $(2f+2)$ new coordinates, call them "parameters" for purposes of distinction, $\xi^1, \dots, \xi^{2f}, \theta, H$. The parameters are not necessarily assumed to be canonical, but the Jacobian of the equations leading from the coordinates to the parameters is assumed to be finite and different from zero, H has the meaning given in Eq. (2.7), and θ satisfies Eq. (2.3). A $2f$ -dimensional hypersurface S , $\theta = \text{const.}$, $H = 0$, will intersect all mechanical trajectories. The desired ensemble density m will have such a significance that an integral extended over part of the hypersurface just introduced,

$$I = \int_{H=0, \theta=\theta_1}^{\theta=\theta_2} m d\xi^1 \dots d\xi^{2f}, \quad (3.2)$$

will equal the fraction of trajectories within the ensemble that are intersected by the chosen domain of integration. If extended over the whole hypersurface, its value should be 1. Once having obtained such a density, we shall also want to be able to carry out similar integrals in terms of the original canonical coordinates.

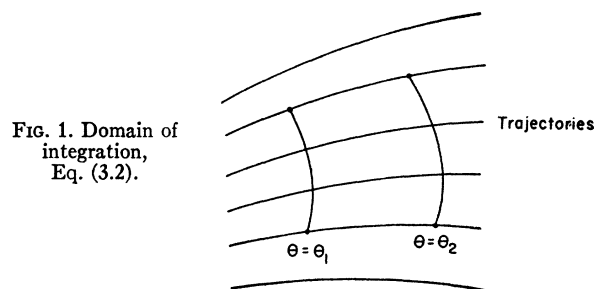


FIG. 1. Domain of integration, Eq. (3.2).

Clearly, the integral I , denoting as it does the ratio between two integrals (both of which may be finite or infinite) must be an invariant. Moreover, its value must be independent of the location of the hypersurface S cutting across the $(2f+1)$ -dimensional hypersurface $H=0$, providing the bounding $(2f-1)$ -dimensional edge intersects the same trajectories. Figure 1 represents this requirement.

It is well known that in the $(2f+1)$ dimensional "reduced space" $H=0$ all these requirements are equivalent to the one that the integrand m be the θ -component of a contravariant vector density \mathbf{C} whose divergence is zero,

$$\partial C^i / \partial \xi^i = 0, \quad i = 1, \dots, 2f+1; \quad \xi^{2f+1} \equiv \theta. \quad (3.3)$$

and which is everywhere parallel to the trajectories. The direction of the trajectories is given in the expanded phase space by the equations of motion. In the reduced space these directions can be represented adequately through projection operations and come out as

$$D^i = \frac{\partial \xi^i}{\partial q_\alpha} \frac{\partial H}{\partial p_\alpha} - \frac{\partial \xi^i}{\partial p_\alpha} \frac{\partial H}{\partial q_\alpha} = (\xi^i, H). \quad (3.4)$$

In the absence of a metric, a direction is given only modulo an arbitrary factor. Let it be noted in passing that the expressions (3.4) transform as the components of a contravariant vector with respect to parameter transformations in the reduced space, i.e., with respect to parameter transformations not involving H .

We may conclude that the vector density \mathbf{C} must be parallel to the vector \mathbf{D} , and that the constant of proportionality must be a scalar density. The principal shortcoming in this preliminary result is that \mathbf{D} also changes when we transform H , even though this transformation has no geometric meaning in the reduced space. The most obvious approach, to construct a vector density in the expanded space and project it into the reduced space, is actually geometrically not feasible. Specifically, although projection leads from a vector or tensor in the higher-dimensional space to a vector or tensor in the lower-dimensional space, the same is not true of densities.

A different approach, seemingly more elaborate, is geometrically sound and leads to the desired vector density in the reduced space. In a space without a Riemannian metric, the completely antisymmetric con-

travariant tensor densities of weight one have the special significance that they can be used to construct invariant integrals of lower dimensionality. In particular, in an n -dimensional space a skewsymmetric contravariant tensor density of rank $(n-m)$, $m \leq n$, permits the construction of an integral over an m -dimensional domain. If we describe this domain by means of a set of m parameters u^i and if the coordinates in the original space are called x^i , the integral

$$I = \frac{1}{(n-m)!} \int \delta_{\alpha_1 \dots \alpha_{n-m} \beta_1 \dots \beta_m} S^{\alpha_1 \dots \alpha_{n-m}} \prod_{i=1}^m \left(\frac{\partial x^{\beta_i}}{\partial u^i} du^i \right), \tag{3.5}$$

$\delta_{1 \dots n} = 1, \quad \delta_{\alpha_1 \alpha_2 \dots} + \delta_{\alpha_2 \alpha_1 \dots} = 0, \text{ etc.},$

is invariant. Now suppose we construct a space of n' dimensions, imbedded in the original n -dimensional space, with $m \leq n' < n$. Then a tensor density can be constructed in the n' -dimensional subspace, such that it leads to exactly the same integral I of Eq. (3.5). If

$$\frac{1}{(n-m)!} \int_u \delta_{\alpha_1 \dots \alpha_{n-m} \beta_1 \dots \beta_m} S^{\alpha_1 \dots \alpha_{n-m}} \prod_{i=1}^m \left(\frac{\partial x^{\beta_i}}{\partial u^i} du^i \right) \equiv \frac{1}{(n'-m)!} \int_u \delta_{r_1 \dots r_{n'-m} s_1 \dots s_m} S^{r_1 \dots r_{n'-m}} \prod_{i=1}^m \left(\frac{\partial v^{s_i}}{\partial u^i} du^i \right). \tag{3.7}$$

Without going into a lower dimensional subspace, we can also gain densities of lower rank by processes of differentiation. The "curl"

$$S^{\alpha_1 \dots \alpha_{n-m+1}, \alpha_{n-m+1}} = f^{\alpha_1 \dots \alpha_{n-m}} \tag{3.8}$$

satisfies the theorem (a generalization of Gauss' and Stokes' theorems) that if the $(m-1)$ parameters w^k describe the edge (closed surface) of the m -dimensional domain (u^i), then

$$\frac{1}{(n-m)!} \int_u \delta_{\alpha_1 \dots \alpha_{n-m} \beta_1 \dots \beta_m} S^{\alpha_1 \dots \alpha_{n-m+1}, \alpha_{n-m+1}} \prod_{i=1}^m \left(\frac{\partial x^{\beta_i}}{\partial u^i} du^i \right) \equiv \frac{1}{(n-m+1)!} \oint_w \delta_{\alpha_1 \dots \alpha_{n-m+1} \beta_1 \dots \beta_{m-1}} S^{\alpha_1 \dots \alpha_{n-m+1}} \prod_{k=1}^{m-1} \left(\frac{\partial x^{\beta_k}}{\partial w^k} dw^k \right). \tag{3.9}$$

Our problem is now to find a vector density in the reduced space which is parallel to a given field of directions and whose divergence vanishes. In the expanded space, we should have to find a contravariant tensor density of rank 2 whose surface integral vanishes for any $2f$ -dimensional closed surface lying wholly in the reduced space. Or, we may require (and this requirement is equivalent) that the divergence of the tensor density, while not zero, vanishes when integrated over a $(2f+1)$ -dimensional surface forming a part of the reduced space.

Let us call the $(2n+2)$ parameters ξ^α , $\alpha=1, \dots, 2f+2$. We shall first show that the contravariant tensor density

$$S^{\alpha\beta} = \mu' (\xi^\alpha, \xi^\beta), \tag{3.10}$$

where μ' is a scalar density, projects into a vector density in the reduced space whose components are proportional to the Poisson brackets (3.4). The proof is almost trivial because we are using the same parameters (except for $\xi^{2f+2} \equiv H$) in both the reduced and the

we call the parameters of the n' -dimensional subspace v^r , we have

$$S^{r_1 \dots r_{n'-m}} = \frac{1}{m!(n-m)!} \delta^{r_1 \dots r_{n'-m} s_1 \dots s_m} \delta_{\alpha_1 \dots \alpha_{n-m} \beta_1 \dots \beta_m} \times \prod_{i=1}^m \left(\frac{\partial x^{\beta_i}}{\partial v^{s_i}} \right) S^{\alpha_1 \dots \alpha_{n-m}}. \tag{3.6}$$

The rank of the new tensor density in the n' -dimensional subspace is lower than the rank of the original density in the expanded space by $(n-n')$. The operation characterized by Eq. (3.6) is, in a certain sense, also a projection in that the new density has fewer components than the original density, and the original density cannot be reconstructed uniquely from the new one. The partial equivalence between the new and the original density is expressed mathematically by the identity

expanded space. In that case, Eq. (3.6) reduces to a rule which requires that we strike out rows and columns in which not at least one index is H.

Next, we shall form the divergence of $S^{\alpha\beta}$, $S^{\alpha\beta, \beta}$, but in a canonical coordinate system, in a canonical coordinate system the tensor density $S^{\alpha\beta}$ assumes the values

$$S^{\alpha\beta} = \begin{pmatrix} 0 & \mu \delta^{\kappa\lambda} \\ -\mu \delta^{\kappa\lambda} & 0 \end{pmatrix}, \quad \kappa, \lambda = 1, \dots, f+1. \tag{3.11}$$

Its divergence, therefore, is

$$S^{\alpha\beta, \beta} = \left(\frac{\partial \mu}{\partial p_\kappa}, -\frac{\partial \mu}{\partial q_\kappa} \right). \tag{3.12}$$

In terms of general parameters, this vector density in the expanded space has the components

$$S^{\alpha\beta, \beta} = (\xi^\alpha, \mu) J, \quad J = \det \left| \frac{\partial (q_\kappa, p_\kappa)}{\partial (\xi^i, \theta, H)} \right|, \quad \mu' = J \mu, \tag{3.13}$$

where J is the Jacobian of the transformation leading from any canonical coordinate system to a parameter system (the Jacobian of a canonical transformation is always 1). The variable μ (without prime) is a scalar. The vector density (3.13) will have vanishing integrals if its H-component is zero, if, in other words,

$$(\mathbf{H}, \mu) = 0. \quad (3.14)$$

This last equation is the invariant formulation of Liouville's theorem in the expanded phase space. It is not a condition satisfied merely by stationary distributions, but a perfectly general condition obeyed by the scalar μ .

With the help of the expression (3.10), where we now replace the density μ' by the expression μJ , we can form the vector density \mathbf{C} in the reduced space that leads to the formation of ensemble densities. We have

$$C^\alpha = (\xi^\alpha, \mathbf{H}) \mu J. \quad (3.15)$$

The integral (3.2) turns into

$$I = \int (\theta, \mathbf{H}) \mu J d\xi^1 \cdots d\xi^{2f}. \quad (3.16)$$

With this expression, it is now possible to describe arbitrary Gibbs ensembles in terms of the expanded canonical formalism. If the function \mathbf{H} is chosen so that it serves as the Hamiltonian, with θ the argument appearing in the canonical equations of motion, then the Poisson bracket in Eq. (3.16) equals 1.

The introduction of the ensemble density operator μ in quantum mechanics is much more conventional and requires no differential-geometrical derivation. Assuming that to each possible Hilbert vector we assign an ensemble probability P_k (where k is a parameter which may either take discrete or continuous values), so that the sum or integral $\sum_k P_k$ equals 1, then, in Dirac's bracket notation, μ is

$$\mu = \sum_k P_k |k\rangle \langle k|. \quad (3.17)$$

The average expectation value of any observable A is given by the expression

$$\langle A \rangle_\mu = \text{tr} \{ \mu A \}, \quad (3.18)$$

and this expression corresponds to the classical integral $\int \mu A dX$.

IV. CANONICAL ENSEMBLES

With an invariant ensemble density, we can now proceed to duplicate the procedures of classical statistical mechanics. The chief difference is that we cannot single out the total energy of the system from among all the other variables of state to construct canonical ensembles. We shall, therefore, generalize the concept of canonical ensemble and call an ensemble canonical whenever it has the same distribution in phase space (or in Hilbert space) that it would assume if its systems were not isolated, but components of much larger

systems (with infinitely many degrees of freedom, in the asymptotic limit) which in turn form a microcanonical ensemble in their phase space.

This generalized definition is based on the physical notion of statistical equilibrium in interaction with a very large constant-temperature bath. Ordinarily, we motivate the assumption of canonicity by saying that a system will assume canonical distribution if it is in interaction with a very large temperature bath and if nothing is known about the combined system except that the total energy lies within a specified narrow interval (E_1, E_2). In that portion of phase space (for the combined system) that is bounded by the energy surfaces E_1 and E_2 the ensemble density is, therefore, assumed to be constant. The canonical distribution is not sensitive to a transition to an energy interval of zero width. To generalize this approach, we must extend it to other variables, and to combinations of other variables. We must assume that it is possible to provide the type of interaction with a large system (bath) in which not only the energy, but also other constants of the motion (for the isolated system) are exchanged, and we must then determine the distribution of the system in which we are interested on the assumption that for the combined system certain variables (e.g., total energy, total linear momentum) are limited to narrow domains, but that within that domain the ensemble density is constant.

This approach will be physically interesting only if the result is insensitive to the detailed structure of the large temperature bath and to the precise nature of the coupling bath and thermodynamic system. That will be the case only if the variables we are interested in are "additive," i.e., if we can naturally construct constants of the motion for the combined system which are sums of two terms, one depending only on the state of the system, the other only on the state of the bath. If that is the case, then all that matters is that the coupling parameter may be permitted to go to zero, and the size of the bath to go to infinity, without destroying the physical meaning of the whole setup. Most of the variables ordinarily of interest will have this property of "additivity." Aside from the standard integrals of the motion, total energy, linear momentum, and angular momentum, the density at a particular location, all sorts of occupation numbers, intensity of radiation in a particular part of the spectrum (all the latter taken at a particular time), and so forth, are additive variables in the sense explained here.

With these assumptions, the system under scrutiny will always assume a distribution in phase space which has the general form

$$\mu = Z^{-1} \exp(-\sum_i \beta_i A_i), \quad (4.1)$$

with

$$Z = \int (\theta, \mathbf{H}) J \exp(-\sum_i \beta_i A_i) d\xi^1 \cdots d\xi^{2f}. \quad (4.2)$$

The variables A_i are those additive constants of the motion for which we have specified the microcanonical distribution in the combined system (or the mean values in the isolated system), Z is the partition function of the ensemble under these conditions, and the β_i are a set of parameters whose values are determined by the required values of the ensemble averages $\langle A_i \rangle_{\mathcal{N}}$. It is, of course, assumed that the quantities A_i are all independent of each other; that none of them may be expressed as a function of the others nor as a function of the other A 's plus the Hamiltonian constraint H .

It is clear that the expression (4.1) is invariant with respect to a transformation that leads from the set A_i to a set A_k' consisting of linear combinations of the A_i . Therefore, an ensemble distribution is relativistically covariant automatically if with any component of a vector or similar covariant among the A_i the other components are included, too.

Equation (4.1) is the natural definition for canonical distributions in the generalized theory. It is interesting to note that with this definition an ensemble may be canonical if its average energy and linear momentum are specified, but also if only its average energy is specified. These two ensembles are by no means identical, but there appears no valid reason for accepting one of them as physically interesting and rejecting the other. It is, of course, true that if we choose as the A_i the four components of the energy-momentum vector, then this choice is relativistically invariant. If in one system we choose just the energy, then in another frame the corresponding choice is a particular linear combination of energy and linear momentum.

V. ENTROPY

All canonical ensembles can be obtained by means of a variational problem with subsidiary conditions. If we define the entropy of an ensemble by means of the integral

$$S = k \int (\theta, H)_{\mu J} \ln(\mu_0/\mu) d\xi^1 \cdots d\xi^{2f} \quad (5.1)$$

(k being the Boltzmann constant and μ_0 an arbitrary constant), then the canonical ensembles are those in which the entropy is maximized under the subsidiary conditions that a set of ensemble averages $\langle A_i \rangle_{\mathcal{N}}$ assumes specified values. The larger the number of these subsidiary conditions, the smaller the value of the entropy. Naturally, we always have the one subsidiary condition (3.1), that $\langle 1 \rangle_{\mathcal{N}}$ equals 1.

The value of the entropy for a canonical ensemble is

$$S = k(\ln \mu_0 + \ln Z + \sum_i \beta_i \langle A_i \rangle_{\mathcal{N}}). \quad (5.2)$$

With the help of this expression, the partition function itself may be represented as a function of purely "thermodynamic" (i.e., macroscopic) variables:

$$Z = \exp \left[\frac{1}{k} (S - \sum_i \beta_i \langle A_i \rangle_{\mathcal{N}} - \ln \mu_0) \right]. \quad (5.3)$$

We shall call the entropy defined by Eq. (5.1) the "ensemble entropy." Because of Liouville's theorem, it is an integral of the motion. In order to find another variable, the "system entropy," which will change its value in the course of the motion, we may proceed as follows. Regardless of whether we have a single system or a canonical or noncanonical ensemble, we define the system entropy S^* by means of Eq. (5.2), where the partition function Z and the multipliers β_i are all determined by the average values $\langle A_i \rangle_{\mathcal{N}}$ on the fictitious assumption of a canonical ensemble. System entropy and ensemble entropy have the same value for a canonical ensemble, for all other cases the system entropy is the larger of the two. That is because for given values of the $\langle A_i \rangle_{\mathcal{N}}$, the canonical distribution is the one that maximizes the ensemble entropy, and the system entropy equals the ensemble entropy of this "corresponding" canonical ensemble. In contrast to the ensemble entropy, the system entropy is not a constant of the motion.

Now consider a set of dynamical variables A_i which are constants of the motion and which depend explicitly on the parameter θ . Naturally, if we construct an ensemble that is canonical with respect to this set of variables, it will remain so, and thus its system entropy as well as its ensemble entropy will have the same value for all values of the parameter θ . But now we can introduce a new set of variables, B_i , by means of the defining equations

$$B_i(\xi^1, \dots, \xi^{2f}, \theta_1) = A_i(\xi^1, \dots, \xi^{2f}, \theta_0) \quad (5.4)$$

and the further requirement that the B_i also be constants of the motion. This requirement is particularly suggested if we should have chosen a set of parameters in which the Hamiltonian constraint H is independent of the parameter θ . Such a choice produces formally a conservative system, and in this case Eq. (5.4) together with the requirement that both the A_i and the B_i be constants of the motion leads automatically to the stronger relationship

$$B_i(\xi^1, \dots, \xi^{2f}, \theta) = A_i(\xi^1, \dots, \xi^{2f}, \theta + \theta_0 - \theta_1). \quad (5.5)$$

In other words, the two sets of variables are quite similar in nature and differ only by a rigid translation of the "time" scale. We shall now consider both the ensemble entropy and the system entropy of an ensemble that is canonical with respect to the A_i , but not with respect to the B_i . Because the ensemble is in general not canonical with respect to the B_i , the system entropy with respect to the B_i will exceed the ensemble entropy. The ensemble entropy, on the other hand, is an intrinsic property of the ensemble and does not depend on whether we consider the A_i or the B_i . It can therefore be asserted that the system entropy with respect to the B_i for our ensemble exceeds the system entropy with respect to the A_i .

Intuitively, the difference between the set of A_i and the set of B_i is that B_i represents the "same" set of

variables, but observed a "time" $(\theta_1 - \theta_0)$ "later"—"earlier" if $(\theta_1 - \theta_0)$ is chosen negative. With respect to the "fixed" observable A_i^0 , i.e., those θ -independent observables which permanently equal $A_i(\theta_0)$, our ensemble is canonical only at the "time" θ_0 . Therefore, at all other "times" the ensemble entropy S is a lower bound for the system entropy S^* , determined with the help of the averages of the A_i^0 .

Let it be noted in passing that this version of the H -theorem holds for the "past" as well as for the "future"—in our theory, all results must be symmetric with respect to the time because nowhere have we introduced into the foundations any element of asymmetry. However, we may introduce this element of asymmetry by focusing our attention on the present and the future only, assuming that any observation requires interaction with our large and bulky measuring devices in such a manner that as a result of this interaction our state of knowledge about the system under observation immediately following the measurement(s) is adequately represented by a canonical ensemble. Repetition of the "same" measurements, i.e., determination of the "fixed" A_i^0 , at some subsequent time will then show an increase of the system entropy. Whether this method of introducing thermodynamic concepts (the Second Law) is completely satisfactory shall not be examined in this paper. At any rate, this approach is the natural analog of that usually taken and, at least, requires no assumptions concerning equilibrium or quasistatic processes.

VI. HEAT FLUX

In order to formulate ordinary thermodynamics, it is necessary to introduce into statistical mechanics the concepts of performance of work and of transfer of heat; these concepts are required to formulate both the first and the second law of thermodynamics. The most satisfactory manner of defining these two quantities in statistical mechanics is first to define the mechanical work and then describe the heat transfer as the difference between the change in energy and the work done on the system. The heat transfer obtained in this manner then leads to an exact differential (the change in entropy) when divided through by the absolute temperature.

In our generalized theory, very similar results may be obtained, but with one significant difference. Instead of the single inexact differential "work," we have such a differential for each of the constants of the motion A_i , describing the adiabatic change of its ensemble average. Starting with a canonical ensemble, we obtain the change in (system) entropy by multiplying the "heat transfer" of each separate variable A_i by its appropriate β_i , and adding them all together. We first define, for a canonical ensemble, the adiabatic change in $\langle A_i \rangle_{Av}$ as

the expression

$$\delta_{ad} \langle A_i \rangle_{Av} = \int \mu \delta A_i dX, \tag{6.1}$$

$$\mu = \frac{1}{Z} e^{-\sum_k \beta_k A_k}, \quad dX = (\theta, H) d\xi^1 \dots d\xi^{2f},$$

which is the analog of the definition given by Schrödinger.⁶ Our parameter system should be chosen so that $\delta[(\theta, H)\mu]$ vanishes. If the total change in $\langle A_i \rangle_{Av}$ is denoted by $\delta \langle A_i \rangle_{Av}$, we find then for the change due to random interaction with the surroundings the expression

$$\delta_Q \langle A_i \rangle_{Av} = \delta \langle A_i \rangle_{Av} - \delta_{ad} \langle A_i \rangle_{Av} = \int A_i \delta \mu dX. \tag{6.2}$$

These expressions are the analogs of the heat transfer. Finally, if we multiply each of these expressions by its β_i and add, we get

$$\begin{aligned} \sum_i \beta_i \delta_Q \langle A_i \rangle_{Av} &= \sum_i \int \beta_i A_i \delta \mu dX \\ &= \delta \int (\sum_i \beta_i A_i) \mu dX - \int \mu \delta (\sum_i \beta_i A_i) dX \\ &= \delta (\sum_i \beta_i \langle A_i \rangle_{Av}) + \delta \ln Z = k^{-1} \delta S^*. \end{aligned} \tag{6.3}$$

This result shows that in a relativistic thermodynamics it is natural to retain the concept of a scalar entropy; but work must be replaced by a set of quantities having similar transformation properties as the chosen A_i , and the same holds for the heat transfer. Thus, if we formulate thermodynamics for a system having a distribution in space (e.g., a finite volume element within a fluid), the heat transfer across its bounding surface should not be a vector, but a tensor or tensor-like formation, which represents not only a flux of energy, but also of linear momentum.

VII. CONCLUSION

In this paper, we have confirmed the intuitive assumption that entropy should be retained as a scalar; but it turns out that the energy, work, and heat concepts must be adjusted in each case to the physical situation encountered and to each other. It is planned to develop the general framework further and to attempt its application to interesting situations far removed from the conventional quasistatic process.

The writer should like to acknowledge with gratitude the stimulation he has received from R. Wildt in many discussions involving the thermodynamics of stellar atmospheres.

⁶ E. Schrödinger, *Statistical Thermodynamics* (Cambridge University Press, Cambridge, 1946), p. 11ff.