Journal of Statistical Physics, Vol. 16, No. 1, 1977

Information-Theoretic Stability and Evolution Criteria in Irreversible Thermodynamics

E. Pfaffelhuber[1]

Received July 15, 1976

Generalized information gains are used to derive stability conditions, for steady states, periodic orbits and invariant sets, and general evolution criteria, both global and local with respect to the time variable, in irreversible thermodynamics. Meixner's passivity condition and the Glansdorff–Prigogine stability and evolution criteria are found to be special cases thereof. The information gain quantities include Kullback's three kinds of divergences, the first two of which are dual to each other and yield criteria which are symmetric in the average densities of the system's extensive variables and the conjugate parameters, but which are nonsymmetric in the irreversible fluxes and forces, while the third one does not involve the entropy function of the system. Furthermore, Renyi's information gain of order α and Csiszar's f-divergence are treated. The latter is used to construct a most general information gain quantity as a Liapunov function and evolution criterion, which, however, for local stability and evolution conditions is still equivalent to the use of the second-order variation of the entropy.

KEY WORDS: Generalized information gain; f-divergence; nonequilibrium thermodynamics; stability theory; Liapunov functions; evolution criteria; generalized canonical distributions; irreversible fluxes and forces.

1. INTRODUCTION AND SUMMARY

In a series of papers[1–11] Schlögl and his collaborators have shown that an information-theoretic quantity, namely the information gain $K(p, p')$ of a probability distribution p with respect to another distribution p', can serve

Supported by the Deutsche Forschungsgemeinschaft.

[1] Institute for Information Sciences, University of Tübingen, Tübingen, Germany (F.R.), and Department of Physics and Astronomy, University of Wyoming, Laramie, Wyoming.

as a Liapunov function for open thermodynamic systems far from equilibrium, and hence as a stability criterion for steady states which generalizes the Glansdorff–Prigogine stability test.[3,4,12,13] Hints were also given concerning the relation between the information gain K and Meixner's passivity conditions[2,7,9,14,15] as well as the Glansdorff–Prigogine evolution criterion.[13] In this work we make use of other information-theoretic quantities, so-called generalized information gains, which have successfully been applied in information theory and statistics. We show that they are also of direct relevance in irreversible thermodynamics, and give rise to a series of stability and evolution criteria.

In Section 2 we introduce three types of information gains K_ν, $\nu = 1, 2, 3$, which have already been used by Kullback,[16] the first one being the conventional one as used by Schlögl. For the case of generalized canonical distributions, K_1 and K_2 are related to each other by means of a Legendre transformation, and K_3 is a symmetrized version of both which has the interesting property that it does not depend directly on any entropy functions. Using these quantities as Liapunov functions, we derive the corresponding three steady-state stability criteria, the first two of which exhibit a complete symmetry between the average densities $\overline{m}^l$ of the system's extensive variables and the corresponding conjugate parameters λ_l, $l = 1,..., f$. The corresponding criterion in terms of the irreversible fluxes and forces, which was already obtained by Schlögl,[3,4] and which generalizes the Glansdorff–Prigogine condition, is rederived. The derivation here is believed to improve the original one in Ref. 3 by taking into account that the irreversible fluxes and forces cannot be directly identified with the conjugate parameters λ_l and time derivatives $\dot{\overline{m}}^l$ of the average densities, respectively, but that beforehand a partial integration must be performed. Only in this way can the familiar expressions for the irreversible fluxes and forces be obtained. The procedure makes explicit that, contrary to the $\overline{m}^l$ and λ_l, the fluxes and forces play a nonsymmetric role in the stability conditions.

Section 3 deals with further extensions of the information gain concept. We use a slight modification of Renyi's information gain of order α,[17] thereby obtaining a one-parameter family of Liapunov functions. It is interesting to note that the corresponding stability criteria do not contain the average densities of the system itself, but rather those of a hypothetical system whose conjugate parameter vector field lies on a straight line in λ-space joining the parameter vector of the system in question and the corresponding steady-state vector. As a further extension we consider Csiszar's f-divergence[18] and form with it the most general information gain like a Liapunov function, which is parametrized by two rather arbitrary functions f and g, with all previously considered information gains being special cases thereof. It is shown that, despite this very general form, for local stability tests, i.e., small deviations from the steady state, the resulting conditions are equivalent, for any admissible choice of f and g, to those obtained from the use of the second-order variation of the entropy[12,13] as a Liapunov function, so that we rediscover the Glansdorff–Prigogine criteria.

Section 4 deals with the stability problem of invariant sets. The basic idea is that the information gain measures some kind of "distance" between two probability distributions, so that it is natural to introduce now as a Liapunov function the minimal "distance" between the system's actual distribution and the invariant set Γ of distributions in question, that is, the minimum of the information gain where the second argument varies over Γ. From the very fact that at the point where the minimum is taken on, the information gain is stationary, we find stability criteria for invariant sets almost identical to the steady-state case. In particular, the Glansdorff–Prigogine steady-state stability condition, with a suitable reinterpretation, holds true also for invariant sets and, in particular, for periodic orbits.

Section 5 establishes a series of local and global temporal evolution criteria. The global form states that a certain two-point function, namely the information gain corresponding to two probability distributions at different times, cannot be negative. If we let these two times approach each other, we obtain a local evolution criterion which implies that the second-order time derivative of the information gain, at the point of coinciding time arguments, must be nonnegative. Various equivalent formulations of these criteria are given for the case of generalized canonical distributions in terms of the average densities $\overline{m^l}$ and conjugate parameters λ_l, as well as in terms of irreversible fluxes and forces. Meixner's passivity condition[14,15] and the Glansdorff–Prigogine evolution criterion[13] result as special cases, as in Refs. 2, 3, 7, and 9. Again, certain criteria exhibit a complete symmetry between the $\overline{m^l}$ and λ_l, whereas a corresponding symmetry does not hold for the irreversible fluxes and forces.

Throughout the paper we do not make any assumptions about the nature of the underlying stochastic processes of the system macroobservables; they may be non-Markovian. We do, however, assume in most cases that the probability distributions at any given time are generalized canonical ones. By applying similar techniques as in Ref. 10, Mori distributions could have been used instead, but would not have led to any novel results in our approach, in which canonical distributions retain all the significant aspects, but avoid the complications, of the Mori distributions. Also it should be noted that the case of moving material could easily be handled by similar techniques as in Ref. 10, and in fact is in principle contained in our approach if the entropy is suitably reinterpreted.

Finally, we point out that the central idea of the paper, namely the use of information-gain-like quantities as Liapunov functions and evolution criteria, is not only most useful in thermodynamics, but has also been applied with considerable success in purely mathematical areas, furnishing a limit theorem for Markov chains,[17,19] a novel proof of the central limit theorem,[20] and as an existence proof for certain optimal hypotheses in statistics.[21]

2. INFORMATION-THEORETIC LIAPUNOV FUNCTIONS

Let i be a discrete label for the microstates of a thermodynamic system, and let p and p' be two probability distributions over these microstates. The quantity

$$K(p, p') = \sum_i p_i \ln(p_i/p_i') \tag{1}$$

is then the conventional information gain as used in Refs. 1–11. Equation (1) is known under various names, such as directed divergence,[16] relative entropy,[22–27] or missing information.[23,25,28] [The apparent contradiction in nomenclature is resolved by the fact that (1) can be interpreted as the gain of information when changing from p' to p, as well as the information still missing before this change is made.] Since the necessary modifications for dealing with continuously varying labels i (as, for example, in classical thermodynamics, where they represent the phase space points) are obvious (replacement of sums by Lebesgue–Stieltjes integrals, and of the ratio p_i/p_i' by the Radon–Nikodym derivative of the two probability measures), here we shall stick for simplicity to the discrete case. Let us define the following quantities:

$$K_1(p, p') = K(p, p') \tag{2}$$

$$K_2(p, p') = K(p', p) = \sum_i p_i' \ln(p_i'/p_i) \tag{3}$$

$$K_3(p, p') = \tfrac{1}{2}[K_1(p, p') + K_2(p, p')] = \tfrac{1}{2} \sum_i (p_i - p_i') \ln(p_i/p_i') \tag{4}$$

Equations (2) and (3) are the two kinds of directed divergence introduced by Kullback[16] (only the first of which has been used so far in connection with steady-state stability tests), and (4) is, apart from the factor 1/2, Kullback's undirected divergence, which is symmetric in p and p' and has already been used by Jeffreys.[29] All the above quantities serve as a measure for discriminating the two probability distributions p and p'. Namely, from the fact that (1) is nonnegative and vanishes if and only if p and p' coincide[1,2,16] it follows immediately that the same is true for (2)–(4). Furthermore, we note that, since (1) is not only concave in p for fixed p',[11] but also concave in p' for fixed p (since the logarithm is a convex function), the quantities (2)–(4) are all concave in each argument separately. Also, an easy calculation shows that if Δ denotes the difference between unprimed and primed quantities, then, neglecting higher than second powers in Δp, we have

$$K_1(p, p') \approx K_2(p, p') \approx K_3(p, p') \approx -T^{(2)}S(p) = \tfrac{1}{2} \sum_i (\Delta p_i)^2/p_i' \qquad (5)$$

where

$$S(p) = \sum_i p_i \log(1/p_i) \qquad (6)$$

is the entropy associated with the distribution p, and $T^{(2)}$ denotes the quadratic term in a Taylor expansion of $S(p)$ around p'.

From the above properties we see that, in the space of distributions over the microstates i, the quantities (2)–(4) may all serve as Liapunov functions to test the stability or attractivity in the large of a probability distribution p' which remains stationary with respect to a given temporal evolution, or of an invariant set of probability distributions.

Let us calculate (2)–(4) for generalized canonical distributions[6–8,30] of thermodynamic systems. Let $M_i = (M_i^1, M_i^2,..., M_i^f)^T$ be an exhaustive set of extensive macroobservables (energy, particle numbers, magnetization, etc.), $m_i = (m_i^1, m_i^2,..., m_i^f)^T$ their densities with respect to volume, $\overline{M} = (\overline{M^1}, \overline{M^2},..., \overline{M^f})^T$ and $\overline{m} = (\overline{m^1}, \overline{m^2},..., \overline{m^f})^T$ the corresponding average values, and $\lambda = (\lambda_1, \lambda_2,..., \lambda_f)$ the conjugate intensive quantities (inverse temperature, negative of chemical potentials divided by temperature, negative of magnetic field strength divided by temperature, etc.). Here T denotes the transpose of a matrix, and we use a system of units where Boltzmann's constant is equal to one. The different components of m_i, $\overline{m}$, and λ may possibly be of different tensorial character with respect to spatial transformations. Instead of considering $\overline{m}$ and λ as elements of the f-dimensional real space, we could as well take $\overline{m}$ as a point on a smooth manifold and λ as the covector field generated by the entropy density (see below) in a natural way. Note that $\overline{m}$ and λ are functions of space and time in general. Let

$$p_i = \exp\left(\phi - \int dv\, \lambda m_i\right) = \exp\left[\int dv\, (\varphi - \lambda m_i)\right] \qquad (7)$$

be the generalized canonical distribution, where ordinary matrix multiplication is assumed throughout, so that λm_i denotes the sum $\sum_{l=1}^{f} \lambda_l m_i^l$. In (7), dv is the (time-independent) volume element, which is macroscopically infinitesimal, but microscopically large enough to contain many "carriers" of the extensive variables, and the integration goes over the whole system. The function

$$\phi = \int dv\, \varphi, \qquad \varphi = \varphi(\lambda) \qquad (8)$$

is the negative of the logarithmic partition function, viz. one of the Massieu functions,[31] and φ is its spatial density. Equation (7) is a certain generalization of Kubo's extensivity ansatz[32,33] to the case of inhomogeneous systems. Assuming a formula analogous to (7) for the primed quantities, one can calculate the information gains (2)–(4) to obtain

$$K_1 = \int dv\, k_1, \qquad k_1 = \Delta\varphi - \Delta\lambda\, \bar{m} = -\Delta s + \lambda'\, \Delta\bar{m} \tag{9}$$

$$K_2 = \int dv\, k_2, \qquad k_2 = -\Delta\varphi + \Delta\lambda\, \bar{m}' = \Delta s - \lambda\, \Delta\bar{m} \tag{10}$$

$$K_3 = \int dv\, k_3, \qquad k_3 = -(1/2)\, \Delta\lambda\, \Delta\bar{m} \tag{11}$$

Here, k_1, k_2, and k_3 are the spatial densities of the information gains (2)–(4); as before, Δ denotes the difference between unprimed and primed quantities; the quantity

$$S = \int dv\, s, \qquad s = s(\bar{m}) \tag{12}$$

from (6) is the entropy corresponding to the distribution (7) and s is its density, with an analogous formula holding for S'. It is noteworthy that K_3 does not depend directly upon any entropy quantities. As such, it deserves special attention if one wishes to pursue the suggestion[34,35] that thermodynamics should be formulated without the latter. We note that $-K_1$ is related to $-K_2$ by means of a Legendre transformation, in the same way as S is related to ϕ. Namely we have

$$s = -\varphi + \lambda\bar{m}, \qquad -k_1 = +k_2 + \Delta\lambda\, \Delta\bar{m}$$

Since different values of λ or $\bar{m}$ correspond to different probability distributions, the nonnegative quantities (9)–(11) vanish only for $\lambda = \lambda'$ and $\bar{m} = \bar{m}'$. Thus, they may all serve as Liapunov functions to test the stability in the large of a stationary distribution p' with corresponding values λ' and $\bar{m}'$. In addition, we note that K_1 is concave in $\bar{m}$ and radially increasing with respect to λ.[11] Similar considerations show that K_2 is concave in λ and radially increasing in $\bar{m}$.

The Liapunov stability test requires the knowledge of the time derivatives of the Liapunov function along the actual trajectory. To this end let us derive the first-order variations of (9)–(11). Since

$$\delta S = \int dv\, \delta s, \qquad \delta s = \lambda\, \delta\bar{m} \tag{13}$$

$$\delta\phi = \int dv\, d\varphi, \qquad \delta\varphi = \delta\lambda\, \bar{m} \tag{14}$$

we find

$$-\delta K_1 = -\int dv\, \delta k_1, \qquad -\delta k_1 = \Delta\lambda\, \delta\bar{m} - \delta\lambda'\, \Delta\bar{m}$$
$$= \Delta\lambda\, \delta(\Delta\bar{m}) - \delta\lambda'\, \Delta\bar{m} + \Delta\lambda\, \delta\bar{m}' \tag{15}$$

$$-\delta K_2 = -\int dv\, \delta k_2, \qquad -\delta k_2 = \delta\lambda\, \Delta\bar{m} - \Delta\lambda\, \delta\bar{m}'$$
$$= \delta(\Delta\lambda)\, \Delta\bar{m} - \Delta\lambda\, \delta\bar{m}' + \delta\lambda' \Delta\bar{m} \tag{16}$$

$$-\delta K_3 = -\int dv\, \delta k_3, \qquad -\delta k_3 = \tfrac{1}{2}\delta(\Delta\lambda)\, \Delta\bar{m} + \tfrac{1}{2}\Delta\lambda\, \delta(\Delta\bar{m})$$
$$= \tfrac{1}{2}\delta(\Delta\lambda\, \Delta\bar{m}) \tag{17}$$

Consequently, if λ' and $\bar{m}'$ are stationary values and δ is now interpreted as a temporal variation, $\delta(\cdots) = \delta t(\partial/\partial t)(\cdots)$, we have

$$-\dot{K}_1 = -\int dv\, \dot{k}_1, \qquad -\dot{k}_1 = \Delta\lambda\, \dot{\bar{m}} = \Delta\lambda\, (\Delta\bar{m})^{\cdot} \tag{18}$$

$$-\dot{K}_2 = -\int dv\, \dot{k}_2, \qquad -\dot{k}_2 = \dot{\lambda}\, \Delta\bar{m} = (\Delta\lambda)^{\cdot}\, \Delta\bar{m} \tag{19}$$

$$-\dot{K}_3 = -\int dv\, \dot{k}_3, \qquad -\dot{k}_3 = \tfrac{1}{2}(\dot{\lambda}\, \Delta\bar{m} + \Delta\lambda\, \dot{\bar{m}}) = \tfrac{1}{2}(\Delta\lambda\, \Delta\bar{m})^{\cdot} \tag{20}$$

where the dot denotes the (phenomenological) time derivative. Since a negative-semidefinite or definite time derivative of a Liapunov function guarantees stability or asymptotic stability, respectively, we conclude that a thermodynamic steady state characterized by λ' and $\bar{m}'$ is stable or asymptotically stable whenever one of the quantities $-\dot{K}_\nu$ from (18)–(20) is positive semidefinite or definite, respectively. In the case of asymptotic stability, the domain of attraction is certainly not smaller than the largest possible region described by $K_\nu \leqslant C$ (for suitable constant C) in the interior of which $-\dot{K}_\nu$ is positive definite. As far as local stability tests (i.e., the case of small deviations of unprimed and primed quantities) are concerned, all three Liapunov functions (9)–(11) are equivalent because of (5). Consequently, they will all result in the well-known Glansdorff–Prigogine criterion[12,13] since K_1 does, as will be shown later (see also Refs. 3 and 4).

We point out that the right-hand sides of formulas (15) and (16) and of Eqs. (18) and (19) can be obtained from each other by interchanging the average densities $\bar{m}$ and $\bar{m}'$ and the corresponding conjugate parameters λ and λ'. This symmetry between $\bar{m}$ and λ, which is also exhibited in (11), (17), and (20), could not have been detected without the use of information gain functions different from (1).

In addition, the following is noteworthy. Assume that for finite times the system's actual probability distribution, say p^{act}, is not necessarily of the canonical form (7), and let

$$\bar{m} = \sum_i p_i^{\mathrm{act}} m_i$$

be the associated average value. Assume further that, if p denotes the associated canonical distribution corresponding to $\bar{m}$, then, along the system's actual trajectory, the inequality

$$\dot{K}(p^{\mathrm{act}}, p) \leqslant 0 \tag{21}$$

holds (note that both p^{act} and p depend on time). As K from (1) measures, in a certain way, the deviation between its two arguments, (21) means that p approaches p^{act} faster than the latter quantity can "move away." Under these circumstances (with p' still assumed to be of the form of a canonical distribution) the above stability criteria involving K_1 are again valid.

To see this, we start from the identity

$$K_1(p^{\mathrm{act}}, p') = K_1(p, p') + K(p^{\mathrm{act}}, p) + \sum_i (p_i^{\mathrm{act}} - p_i) \ln(p_i/p_i') \tag{22}$$

Because $\ln(p_i/p_i')$ is linear in m_i, and since the average value of the latter quantity is the same under both distributions p^{act} and p, according to our assumption, the sum in (22) drops out. Now the nonnegative (positive) definiteness of $-\dot{K}_1(p^{\mathrm{act}}, p')$ is clearly a sufficient criterion for (asymptotic) stability of the state characterized by the distribution p', so that, by (21), the same is true for $-\dot{K}_1(p, p')$, as was to be shown.

This result seems important since it allows for systems for which a conjugate parameter field does not really exist (e.g., the local temperature need not be defined). Instead, the above stability criterion uses the parameter field that would exist if the distribution were a canonical one.

To relate the quantities on the right-hand side of (18) to the thermodynamic fluxes and forces, we observe, first of all, that the densities $\overline{m^l}$ satisfy an equation of continuity,

$$\dot{\overline{m^l}} = \sigma_{m^l} - \operatorname{div} j_{m^l} \tag{23}$$

where σ_{m^l} is the source term and j_{m^l} (which is a vector in ordinary three-space, provided the lth extensive variable is a scalar, and, generally speaking, is of tensorial rank one higher than the corresponding extensive variable) represents the flow term. Equation (23) will be simply written as

$$m = \sigma_m - \operatorname{div} j_m \tag{24}$$

where σ_m, j_m, and $\operatorname{div} j_m$ are (column) vectors with components σ_{m^l}, j_{m^l}, and

$\operatorname{div} j_{m^l}$, respectively. From (13) we have, interpreting δ again as a temporal variation, that the time change of the entropy density is given by

$$\dot{s} = \lambda \dot{\bar{m}} \tag{25}$$

Using (24), we may write, after performing a partial integration,

$$\dot{s} = \sigma_s - \operatorname{div} j_s \tag{26}$$

where

$$\sigma_s = \lambda \sigma_m + (\operatorname{grad} \lambda) j_m \tag{27}$$

is the entropy production [with grad λ being considered a (row) vector with components grad λ_l], and

$$j_s = \lambda \cdot j_m \tag{28}$$

is the entropy flow. Obviously, the entropy production (27) may be written in the familiar form

$$\sigma_s = X \cdot J \tag{29}$$

where the (row) vector X of the thermodynamic forces X_l is linear in λ and grad λ,

$$X = \lambda \cdot A_1 + (\operatorname{grad} \lambda) \cdot A_2 \tag{30}$$

and the (column) vector J of the thermodynamic fluxes J^l is linear in σ_m and j_m,

$$J = B_1 \cdot \sigma_m + B_2 \cdot j_m \tag{31}$$

and A_1, A_2, B_1, and B_2 are constant matrices. (Of course, they are conveniently chosen such that no mixture of tensorial quantities of different rank will occur in any of the components of X and J.)

We remark that we do not follow here the suggestion in Ref. 3 to identify the time derivatives $\dot{\bar{m^l}}$ of the average densities directly with the irreversible fluxes, and the components of λ with the forces. In fact, from formulas (24), (26), and (29)–(31) one can see that only for spatially homogeneous systems is one allowed to make this identification, because then grad λ vanishes. In all other cases, however, the gradients of the λ_l (inverse temperature, chemical potentials, etc.) cannot be neglected and are clearly most significant components of the irreversible forces. Also, only with the above formulation does one get the flows j_{m^l} (heat flow, particle flow, etc.) as components of J, as it should be.

The physical conservation laws will in general restrict σ_m to lie in a linear subspace of the whole f-dimensional real space, which means that some components of σ_m can be expressed in terms of the others. For example, the conservation of atomic species in chemical reactions leads to an equation of the form $\beta \cdot \sigma_m = 0$, where the components of the matrix β are the numbers of atoms of the various species which compose the different molecules taking part in the reactions.[36] Similarly, energy or charge conservation demands that the sum of all the energy, or charge, source terms vanishes, so that again one of the source terms can be expressed in terms of the others. Also, had we chosen a set of extensive observables such that the average spatial densities $\overline{m}^l$ were functionally dependent (as is the case if we take, e.g., all concentrations plus the overall mass density), we would have a corresponding functional dependence between the components of σ_m and j_m. These types of functional dependence will clearly not invalidate (30) and (31) since the only requirement, which follows from (27) and (29), is that

$$A_i B_j = 1 \cdot \delta_{ij}; \qquad i, j = 1, 2 \tag{32}$$

on the subspace of all physically admissible vectors σ_m and j_m, where 1 is the unit matrix and δ_{ij} is the Kronecker symbol. [We remark that, if (32) were generally valid, as it might be in some special cases, it would allow us to write the entropy flow term (28) as $j_s = XB_1A_2J$.]

Since the last expression for $-\dot{k}_1$ in (18) is obtainable from (25) by replacing λ and $\overline{m}$ by the corresponding differences $\Delta\lambda$ and $\Delta\overline{m}$, and because of the linear relations (30) and (31), it is obvious that we can write down an equation of continuity for k_1 analogous to (26), (28), and (29) with X and J replaced by ΔX and ΔJ, and a change in sign. Thus

$$\dot{k}_1 = \sigma_{k_1} - \operatorname{div} j_{k_1} \tag{33}$$

where

$$-\sigma_{k_1} = \Delta X \, \Delta J \tag{34}$$

$$-j_{k_1} = \Delta\lambda \, \Delta j_m \tag{35}$$

Thus, if

$$-j_{k_1} = \Delta\lambda \, \Delta j_m = 0 \quad \text{on the boundary} \quad \partial\Sigma \tag{36}$$

of the system (which is true if, on each part of $\partial\Sigma$, λ or j_m—or, more generally, some of the components of λ and those components of j_m with different indices—have preassigned, time-independent values) then, by performing a partial integration in (18), we find

$$-\dot{K}_1 = \int dv \, \Delta X \, \Delta J \tag{37}$$

Hence, a sufficient criterion for the stability, or asymptotic stability, of the steady state described by λ', $\overline{m}'$, X', and J' is that (37) is nonnegative, or positive, definite, respectively. This result has been obtained already in Ref. 3 with, however, a slightly different interpretation of X and J, as was mentioned before. For small deviations from the steady state (37) is the well known Glansdorff–Prigogine criterion.[12,13]

If we had used the first expression for $-\dot{k}_1$ in (18) in the above derivation, then by similar techniques we would have found instead of (33)–(35) that

$$\dot{k}_1 = \hat{\sigma}_{k_1} - \operatorname{div} \hat{j}_{k_1} \tag{38}$$

$$-\hat{\sigma}_{k_1} = \Delta X J \tag{39}$$

$$-\hat{j}_{k_1} = \Delta\lambda \cdot j_m \tag{40}$$

so that

$$-\dot{K}_1 = \int dv \, \Delta X J \tag{41}$$

provided that

$$-\hat{j}_{k_1} = \Delta\lambda \cdot j_m = 0 \quad \text{on} \quad \partial\Sigma \tag{42}$$

Hence, under (42), the nonnegative (positive) definiteness of (41) is also a sufficient criterion for (asymptotic) stability, which, for small ΔX, becomes identical to another local steady-state stability condition given by Glansdorff and Prigogine.[13,14] Of course, (33) and (38) are equivalent, since (35) and (40) differ by $\Delta\lambda \cdot j_m$, while (34) and (39) differ by the divergence of this vector field.

3. GENERALIZED INFORMATION GAINS

The concept of the gain of information has been extended by Renyi,[17] who has also given an axiomatic characterization resulting in a one-parameter family of quantities, the so-called information gains of order α, given by

$$(\alpha - 1)^{-1} \log \sum_i p_i(p_i/p_i')^{\alpha-1} \tag{43}$$

Since (43) is nonnegative for $\alpha > 0$, but nonpositive for $\alpha < 0$, it is more advantageous for our purposes to divide (43) by α, so that the resulting expression will be always nonnegative. We obtain the quantity

$$K^{(\alpha)}(p, p') = [\alpha(\alpha - 1)]^{-1} \log \sum_i p_i(p_i/p_i')^{\alpha-1} \tag{44}$$

where α is an arbitrary real number, and, as is easy to check, the limiting cases $\alpha \to 1$ and $\alpha \to 0$ yield, respectively, the quantities (2) and (3). Since (44) is concave in each argument separately and vanishes if and only if the two distributions p and p' coincide, it may again serve as a Liapunov function, parametrized by α. Because

$$K^{(\alpha)}(p, p') = K^{(1-\alpha)}(p', p)$$

there is no need to generate new quantities by interchanging p and p'. The analog of (5) reads

$$K^{(\alpha)}(p, p') \approx -T^{(2)}S(p) = (1/2) \sum_i (\Delta p_i)^2/p_i'$$

and we see that, as far as local stability tests are concerned, the functions (44), for different α, are all equivalent and lead again to the Glansdorff–Prigogine condition.

For canonical distributions (7) we obtain

$$K^{(\alpha)} = \int dv\, k^{(\alpha)}, \qquad k^{(\alpha)} = [\alpha(1-\alpha)]^{-1}[\varphi'' - \alpha\varphi - (1-\alpha)\varphi']$$

where $\varphi'' = \varphi(\lambda'')$ and

$$\lambda'' = \alpha\lambda + (1-\alpha)\lambda' \tag{45}$$

From (14) we find

$$\begin{aligned} -\dot{K}^{(\alpha)} &= (1-\alpha)^{-1} \int dv\, \dot{\lambda}(\bar{m} - \bar{m}'') \\ &= (1-\alpha)^{-2} \int dv\, (\lambda - \lambda'')\cdot(\bar{m} - \bar{m}'') \end{aligned} \tag{46}$$

where $\bar{m}''$ corresponds to the parameter λ'' from (45), so that we obtain a one-parameter family of stability conditions. An interesting case is $\alpha = 1/2$, with

$$K^{(1/2)}(p, p') = -4 \log \sum_i (p_i p_i')^{1/2}$$

being related to the Hellinger integral, and

$$k^{(1/2)} = -4\left[\varphi\left(\frac{\lambda + \lambda'}{2}\right) - \frac{1}{2}\varphi(\lambda) - \frac{1}{2}\varphi(\lambda')\right]$$

It should be noted that the above stability criteria—contrary to the previous ones—involve conditions on the double-primed quantities which correspond to a hypothetical system whose conjugate parameter vector (45) lies, in λ-space, on a straight line connecting the points λ and λ' and in between these points only for $0 < \alpha < 1$.

A further extension of the concept of the gain of information is Csiszar's f-divergence,[18,19]

$$K_f(p, p') = \sum_i p_i f(p_i'/p_i) \tag{47}$$

Here, f is an arbitrary concave function, defined for nonnegative arguments; by subtracting a suitable constant from f we can always achieve $f(1) = 0$, so that (47) vanishes for $p = p'$. It is known[18] that, if f is strictly convex at the point 1, then (47) vanishes only for $p = p'$. Thus, it is positive definite, and may again serve as a Liapunov function. [Note that the choices $f(\xi) = \log(1/\xi)$, $f(\xi) = \xi \log \xi$, and $f(\xi) = (\xi - 1) \log \xi$ yield, respectively, the expressions (2)–(4).] Again there is no need to generate new quantities by interchanging p and p' in (47) since $K_f(p', p) = K_{\hat{f}}(p, p')$, where $\hat{f}(\xi) = \xi \cdot f(\xi^{-1})$ is convex if f is. For canonical distributions (7) we obtain formally

$$K_f = e^{\phi} f\left(\exp\left[-\Delta\phi - \int dv\, \Delta\lambda\, \delta/\delta\lambda\right]\right) e^{-\phi}$$

where $\delta/\delta\lambda$ is a (column) vector with components $\delta/\delta\lambda_l$, and these functional derivatives act solely on the right factor $e^{-\phi}$, which is a functional of λ according to (8).

Since a monotonically increasing function g (vanishing at the origin) of a Liapunov function is again a Liapunov function, we may use as such the quantity

$$V(p, p') = g(K_f(p, p')) \tag{48}$$

For the choice $f(\xi) = \{\text{sign}[\alpha(\alpha - 1)]\}(\xi^{1-\alpha} - 1)$ and $g(\xi) = [\alpha(\alpha - 1)]^1 \times \log\{\xi/\text{sign}[\alpha(\alpha - 1)] + 1\}$ we obtain again the quantity (44). It is interesting that, despite the general nature of (48), which contains two rather arbitrary functions f and g, for local stability tests $V(p, p')$ is nevertheless essentially equivalent to the simple Liapunov function $-T^{(2)}S(p)$ from (5). Namely, if $g^{(\nu)}(0)$ is the first nonvanishing derivative of g at the origin, then we have for small Δp the asymptotic relation

$$V(p, p') = \text{const} \times [-T^{(2)}S(p)]^{\nu} \tag{49}$$

where the (positive) constant is $g^{(\nu)}(0)[f''(1)/2]^{\nu}$, as follows immediately by expanding (48) in terms of powers of Δp.

4. STABILITY OF INVARIANT SETS AND PERIODIC ORBITS

Since a positive-definite function $V(p, p')$ measures, in some sense, the "distance" between the points p and p', it seems natural, with respect to the problem of stability and attractivity of an invariant set Γ (i.e., a set which, once entered, is never left again by the system's temporal trajectory) to use as a Liapunov function the "distance" between the point p and the set Γ, i.e., the function

$$V(p, \Gamma) = \min_{p' \in \Gamma} V(p, p') = V(p, p'(p)) \tag{50}$$

assuming that the minimum exists for all points p (note that Γ may be assumed to be closed[37]) and is taken on at the point $p' = p'(p)$, say.

We have the following statement, which apparently has not been used so far in general stability theory. A sufficient criterion for the (asymptotic) stability of the set Γ is that the time change of $V(p, p'(p))$, which is caused solely by the temporal change of the first argument in V, is nonpositive (negative) definite. That is to say, we must have

$$\sum_i \dot{p}_i \, \partial V(p, p'(p))/\partial p_i \underset{(<)}{\leqslant} 0 \tag{51}$$

where $\partial V/\partial p_i$ means differentiation of V with respect to its first argument only. To prove this, it is sufficient to show that (50) is monotonically decreasing unless $p \in \Gamma$. Letting $\gamma > 0$, we have

$$\begin{aligned} &V(t + \gamma) - V(t) \\ &\quad \equiv V\big(p(t + \gamma), p'(p(t + \gamma))\big) - V\big(p(t), p'(p(t))\big) \\ &\quad = \big[V\big(p(t + \gamma), p'(p(t + \gamma))\big) - V\big(p(t + \gamma), p'(p(t))\big)\big] \\ &\qquad + \big[V\big(p(t + \gamma), p'(p(t))\big) - V\big(p(t), p'(p(t))\big)\big] \end{aligned}$$

By the definition of the function $p'(\cdot)$, the first bracket on the right-hand side is always $\leqslant 0$. It is thus sufficient to show that the second term is always nonpositive (negative), which in fact is guaranteed by our assumption (51). Note that the differentiability of $p'(p)$ with respect to p was not needed in the proof, and in fact will not hold generally if Γ is not convex. A simple example for this case is sketched in Fig. 1. Here, if p varies along the line L, we will obviously have $p'(p) = a$ for p varying between a and b, and $p'(p) = c$ for p

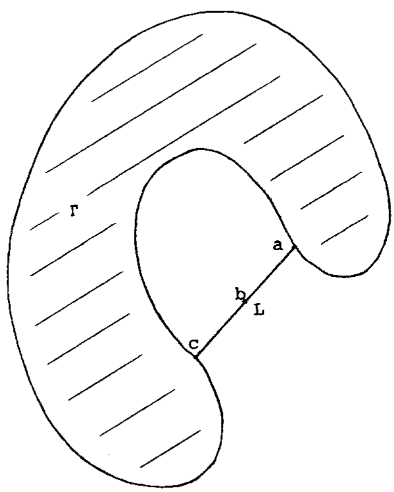

Fig. 1

lying between b and c, where b is an appropriately chosen point on L. Thus, $p'(p)$ is a noncontinuous, and hence nondifferentiable, function here.

Similarly as for steady states, one finds that the domain of attraction, in the case of asymptotic stability, includes all those points p for which (50) is bounded by a suitable constant C, with (51) being valid.

The important point in the above criterion is to realize the very fact that only the temporal change of $V(p, p')$ as caused by its dependence upon the first argument p is involved. Hence we end up with essentially the same conditions as for stationary p'. Consequently, if we use for V one of the functions (2)–(4), (44), (47), or (48), we reocver exactly the same stability conditions as in the steady-state case, given in the previous two sections, except that now the primed quantities must be interpreted as those points where V, as a function of its second argument with the first one held fixed, takes on its minimum according to (50). Explicitly we find from (18)–(20) and (46) that a sufficient criterion for (asymptotic) stability of the set Γ is that one of the expressions

$$\int dv\, \Delta\lambda\, \dot{\bar{m}}, \qquad \int dv\, \dot{\lambda}\, \Delta\bar{m}, \tag{52}$$

$$\tfrac{1}{2}\int dv\, (\dot{\lambda}\, \Delta\bar{m} + \Delta\lambda\, m), \qquad (1-\alpha)^{-1}\int dv\, \dot{\lambda}(\bar{m} - \bar{m}'')$$

is nonnegative (positive) definite. Also, if the boundary condition (42) holds, the first expression in (52) is the same as

$$\int dv\, \Delta X\, J \tag{53}$$

as was shown already in Section 2. Consequently, for the stability criterion in local form (i.e., for small deviations from the set Γ, the latter being of course of finite extension) the Glansdorff–Prigogine criterion in the form (53) with small ΔX is also valid for invariant sets, and hence in particular for periodic orbits. We assumed here that the probability distributions are again canonical ones, which, however, seems less likely to hold than in the case of systems tending toward a stationary state.

The above results can also be found by a more explicit calculation of the minimum of V according to (50), provided this minimum can be obtained by setting the derivative of V with respect to the primed variables equal to zero. As pointed out above, this will not be true in general unless we suitably restrict the domain of the system variables under consideration, namely to an appropriate neighborhood of Γ, as we shall do here. For convenience we take for V the quantity (2) and consider the case of a periodic orbit, which in λ-space is the set

$$\Gamma = \{\lambda'(\sigma):\quad 0 \leqslant \sigma < \tau\} \tag{54}$$

where τ is the period. [The other information gain functions introduced previously could be used as well; also, more general sets Γ could be handled in this manner provided their boundary is sufficiently smooth so that the method of Lagrange parameters works when evaluating the minimum (50); finally, even the above differentiability assumption could be dropped by developing a Kuhn–Tucker-type argument.] Expressing p by means of the corresponding mean value vector field $\bar{m}$, and p' by the conjugate parameter field λ', we have $V = K_1$ as a functional of $\bar{m}$ and λ'.

To evaluate the minimum (50) when λ' varies over Γ, we put $(d/d\sigma)K_1 = 0$, i.e.,

$$\int dv\, \mathring{\lambda}'\, \delta K_1/\delta\lambda' = 0$$

where the dot indicates the derivative with respect to σ. Using (15), we find

$$\int dv\, \mathring{\lambda}'\, \Delta\bar{m} = 0 \tag{55}$$

Let σ_0 be the solution of (55), and let us denote $\lambda'(\sigma_0)$ again by λ', so that the Liapunov function (50), in $\bar{m}$-space, is given by (9). Note that $\lambda' = \lambda'(\sigma_0)$ is the point on the periodic orbit in λ-space that is closest, in the sense of a minimal value of K_1, to the (time-varying) point λ that characterizes the state of the system under consideration. The (asymptotic) stability of (54) is now guaranteed if the time change of $-V = -K_1$ is nonnegative (positive) definite. Using (15), we see that

$$-\dot{K}_1 = \int dv\, (\Delta\lambda\, \dot{\bar{m}} - \sigma_0\mathring{\lambda}'\, \Delta\bar{m})$$

which, by (55), reduces to

$$-\dot{K}_1 = \int dv\, \Delta\lambda\, \dot{\bar{m}}$$

and coincides with the first quantity listed in (52).

5. INFORMATION-THEORETIC EVOLUTION CRITERIA

In a very general sense the equation

$$V(p(t_1), p(t_0)) \geqslant 0 \tag{56}$$

where V is a positive-definite function, represents an evolution criterion, since it implies that the system's temporal development from time t_0 with corresponding probability distribution $p(t_0)$ to time t_1 with corresponding distribution $p(t_1)$ (forward or backward in time according to whether t_1 is larger or smaller than t_0) can only be such that (56) holds. Rewriting (56) as

$$-\int_{t_0}^{t_1} dt\, (d/dt) V(p(t), p(t_0)) \leqslant 0 \tag{57}$$

and taking for V one of the information gain functions (2)–(4), (44), (47), or (48), we end up with a whole series of evolution criteria. We note that the dual of (57),

$$\int_{t_0}^{t_1} dt\, (d/dt) V(p(t_1), p(t)) \leqslant 0$$

will not bring anything new here since it essentially corresponds to an interchange of p and p' in the information gain functions, which gives rise simply to another kind of such function. If we take $V = K_1$ from (2) and use (18) with the primed quantities now being interpreted as the values of the corresponding variables at time t_0, we find from (57)

$$\int_{t_0}^{t_1} dt \int dv\, \Delta\lambda\, (\Delta\bar{m})^{\cdot} \equiv \int_{t_0}^{t_1} dt \int dv\, \Delta\lambda\, \dot{\bar{m}} \leqslant 0 \tag{58}$$

where here Δ denotes the difference between the variables at time t and at time t_0. Equation (58) can be integrated to give, using (13) and (25) and replacing t_1 again by t,

$$\Delta S \leqslant \int dv\, \lambda'\, \Delta\bar{m} \tag{59}$$

If, in addition, the boundary condition (36) holds, then, by similar arguments as in Section 2, (58) can be written as

$$\int_{t_0}^{t_1} dt \int dv\, \Delta X\, \Delta J \leqslant 0 \tag{60}$$

Similarly, if the boundary condition (42) holds, then (58) is equivalent to

$$\int_{t_0}^{t_1} dt \int dv\, \Delta X J \leqslant 0 \tag{61}$$

Expression (61) is one of Meixner's passivity conditions[14,15] and has been derived already in Refs. 2 and 7, where, however, again the interpretation of the quantities X and J is slightly different from ours. We remark that the boundary conditions (36) and (42) are both valid whenever

$$\Delta\lambda = 0 \quad \text{on} \quad \partial\Sigma$$

which means stationary boundary conditions for the conjugate parameter vectors, since λ' is a stationary quantity. Taking for V the quantities (3), (4), and (44), respectively, one derives in a similar way that

$$\int_{t_0}^{t_1} dt \int dv\,(\Delta\lambda)^{\cdot}\,\Delta\bar{m} \equiv \int_{t_0}^{t_1} dt \int dv\,\dot{\lambda}\,\Delta\bar{m} \leqslant 0 \tag{62}$$

$$\int_{t_0}^{t_1} dt \int dv\,(\Delta\lambda\,\dot{\bar{m}} + \dot{\lambda}\,\Delta\bar{m}) \equiv \int_{t_0}^{t_1} dt \int dv\,(\Delta\lambda\,\Delta\bar{m})^{\cdot} \leqslant 0 \tag{63}$$

and

$$(1-\alpha)^{-1} \int_{t_0}^{t_1} dt \int dv\,\dot{\lambda}(\bar{m} - \bar{m}'') \leqslant 0 \tag{64}$$

where $\bar{m}''$ corresponds again to the parameter vector λ'' from (45), with λ and λ' being the parameters at times t and t_0, respectively. We note that (58) and (62) exhibit again a complete symmetry between the average densities and the conjugate parameters; this symmetry also shows up in (63). Using (14), (62) can be integrated to give

$$\Delta\phi \leqslant \int dv\,\Delta\lambda\,\bar{m}' \tag{65}$$

as a dual of (59). The inequality (63) is particularly interesting because of its simplicity. Integrating (63) gives the simple evolution criterion

$$\int dv\,\Delta\lambda\,\Delta\bar{m} \leqslant 0 \tag{66}$$

which does not contain any thermodynamic functions but solely the vector fields $\bar{m}$ and λ. We remark that, because of (9)–(11), the criteria (59), (65), and (66) also could have been obtained directly from (56) by taking for V the quantities (2)–(4).

As has been pointed out in Section 3, the different information gain functions (2)–(4), (44), (47), and (48) are all essentially equivalent locally, i.e., for small values of $\Delta p = p - p'$. Consequently, the evolution criteria generated by them will in turn all be equivalent for t_1 close to t_0, i.e., for small time intervals, because then also $p(t_1)$ will be close to $p(t_0)$. Hence, in order to derive this criterion, we may pick the most convenient information gain function, which is (4) in this case, and which has given rise to (66). Dividing the latter inequality by $(\Delta t)^2$, where $\Delta t = t_1 - t_0$, and letting $\Delta t \to 0$, we find the evolution criterion

$$\int dv\,\dot{\lambda}\dot{\bar{m}} \leqslant 0 \tag{67}$$

We remark that (67) is equivalent to

$$(d^2/dt^2)V(p(t), p(t_0))|_{t=t_0} \geqslant 0 \tag{68}$$

where V is any one of the previously used information gains (2)–(4), (44), (47), or (48), provided that g in (48) has a nonvanishing first derivative at the origin. Namely, from (49) with $\nu = 1$ and (5) it follows that, for t close to t_0, we have

$$V(p(t), p(t_0)) \approx \text{const} \times K_3[p(t), p(t_0)]; \qquad \text{const} > 0$$

while (11) shows that then

$$K_3 \approx -(1/2)(\Delta t)^2 \int dv\, \dot{\lambda}\dot{\bar{m}}$$

which proves (68).

The relation (68) can also be proved directly from (56), which obviously implies that the first nonvanishing term in a Taylor expansion of $V(p(t), p(t_0))$ around $t = t_0$ must be nonnegative. Assuming that already the first term in this expression is different from zero, this implies that

$$(t - t_0)(d/dt)V(p(t), p(t_0))|_{t=t_0} > 0 \qquad \text{for small} \quad t - t_0 \neq 0 \tag{69}$$

Now, (69) is impossible because $t - t_0$ may have either sign. Consequently, the linear term of the Taylor expansion of (56) vanishes, and so we conclude that (68) is valid. The relations (56) and (68) are evolution criteria of a very general nature. They make use of an information gain quantity V which may be called a two-point function with respect to the time variables.

In the following some equivalent forms of (67) and (68) will be derived. Putting (13) in the form

$$\dot{S} = \int dv\, \dot{s}, \qquad \dot{s} = \lambda\dot{\bar{m}}$$

and writing, in a self-explanatory notation,

$$\ddot{S} = \frac{d\dot{S}}{dt} = \frac{d\lambda}{dt}\dot{S} + \frac{d\dot{\bar{m}}}{dt}\dot{S}$$

where

$$\frac{d\lambda}{dt}\dot{S} = \int dv\, \dot{\lambda}\dot{\bar{m}}, \qquad \frac{dm}{dt}\dot{S} = \int dv\, \lambda\ddot{\bar{m}} \tag{70}$$

we have from (67)

$$(d\lambda/dt)\dot{S} \leqslant 0 \tag{71}$$

This criterion says that the acceleration of the total entropy of the system, caused solely by the time dependence of λ, cannot be positive at any time. Similarly, putting (14) in the form

$$\dot{\phi} = \int dv\, \dot{\varphi}, \qquad \dot{\varphi} = \dot{\lambda}\bar{m}$$

and writing

$$\ddot{\phi} = \frac{d\dot{\phi}}{dt} = \frac{d\lambda}{dt}\dot{\phi} + \frac{dm}{dt}\dot{\phi}$$

with

$$(d\lambda/dt)\dot{\phi} = \int dv\, \ddot{\lambda}\dot{\bar{m}}$$

$$(dm/dt)\dot{\phi} = \int dv\, \dot{\lambda}\ddot{\bar{m}}$$

we find

$$(dm/dt)\dot{\phi} \leqslant 0 \tag{72}$$

which is the dual of (71).

Denoting now by P the total entropy production of the system,

$$P = \int dv\, \sigma_s = \int dv\, XJ$$

and putting

$$\dot{P} = \frac{dX}{dt}P + \frac{dJ}{dt}P$$

where

$$\frac{dX}{dt}P = \int dv\, \dot{X}J, \qquad \frac{dJ}{dt}P = \int dv\, X\dot{J} \tag{73}$$

and assuming that the boundary condition

$$\dot{\lambda}j_m = 0 \quad \text{on} \quad \partial\Sigma \tag{74}$$

is valid, we have, by similar arguments as in Section 2, that the right-hand sides of (70) and (73) coincide, so that

$$(dX/dt)P \leqslant 0 \tag{75}$$

Similarly, if

$$\hat{P} = \int dv\, \Delta X\, \Delta J$$

denotes the so-called excess entropy production, where Δ denotes the deviations from steady-state values, then under the boundary condition

$$\dot{\lambda}\, \Delta j_m = 0 \quad \text{on} \quad \partial\Sigma \tag{76}$$

we find

$$(dX/dt)\hat{P} \leqslant 0 \tag{77}$$

where

$$\frac{dX}{dt}\hat{P} = \int dv\, (\Delta X)^{\cdot}\Delta J = \int dv\, \dot{X}\, \Delta J$$

The relations (56), (58), (59), (62)–(68), (71), and (72) are generally valid evolution criteria. Relations (75) and (77) represent the well-known Glansdorff–Prigogine evolution conditions[13] valid under the boundary conditions (74) and (76), respectively, with (60) and (61) being the corresponding forms for finite time intervals.

ACKNOWLEDGMENTS

The author is most grateful to Prof. W. T. Grandy for helpful and stimulating discussions, and for the hospitality of the Department of Physics and Astronomy, University of Wyoming, where part of this work has been done. He is indebted to the Deutsche Forschungsgemeinschaft for financial support.

REFERENCES

1. F. Schlögl, *Z. Phys.* **191**:81 (1966).
2. F. Schlögl, *Z. Phys.* **195**:98 (1966).
3. F. Schlögl, *Z. Phys.* **243**:303 (1971).
4. F. Schlögl, *Z. Phys.* **248**:446 (1971).
5. F. Schlögl, *Z. Phys.* **249**:1 (1971).
6. F. Schlögl, Zur statistischen Thermodynamik um einen stationäre Zustand, Fortbildungsseminar über Thermodynamik irreversibler Vorgänge, MPI für Biophysikal. Chemie, Göttingen, Germany (1972).
7. F. Schlögl, Statistical Inequalities and Thermodynamics of Nonisolated Systems, Intern. Symp. Foundations of Continuum Thermodynamics, Portugal (1973).
8. F. Schlögl, Reversibility Invariants and Nonlinear Thermodynamics, Intern. School of Quantum Electronics, Italy (1974).
9. R. Bausch, *Z. Phys.* **244**:190 (1971).
10. H. K. Janssen, *Z. Phys.* **253**:176 (1972).
11. R. Bessenrodt, *Z. Phys.* **264**:186 (1973).
12. P. Glansdorff and I. Prigogine, *Physica* **46**:344 (1970).
13. P. Glansdorff and I. Prigogine, *Thermodynamic Theory of Structure, Stability and Fluctuations*, Wiley (1971).
14. J. Meixner, *Z. Phys.* **219**:79 (1969).
15. J. Meixner, *Arch. Rat. Mech. Anal.* **33**:33 (1969).
16. S. Kullback, *Information Theory and Statistics*, Dover (1968).
17. A. Renyi, *Wahrscheinlichkeitsrechnung*, VEB (1971).
18. I. Csiszar, *Stud. Sci. Math. Hung.* **2**:299 (1967).
19. I. Csiszar, *Magyar Tud. Akad. Mat. Kutato Int. Közl.* **8**:85 (1963).
20. J. W. Linnik, *Teor. Werojatn. Prim.* **4**:311 (1959).
21. F. Liese, On the Existence of f-Projections, in *Proc. Coll. Information Theory*, Hungary (1975).
22. M. S. Pinsker, *Information and Information Stability of Random Variables and Processes*, Holden-Day (1964).
23. E. Pfaffelhuber, *Kybernetik* **13**:229 (1973).
24. E. Pfaffelhuber, Finite and Infinite Memory Range Learning Processes in Stationary

and Non-Stationary Environment, in *Proc. IEEE Symp. Adapt. Processes*, California (1973).

25. E. Pfaffelhuber, Information Theory and Learning in Biology, in *Physics and Mathematics of the Nervous System*, M. Conrad, W. Güttinger, and M. Dal Cin, eds., Springer LN in Biomath. (1974), Vol. 4, p. 502.
26. E. Pfaffelhuber, Approximate Source Probabilities, Relative Entropy, and Realistic High Probability Groups, in *Proc. IEEE Intern. Symp. Information Theory*, Israel (1973).
27. E. Pfaffelhuber, Minimax Information Gain and Minimum Discrimination Principle, in *Proc. Coll. Information Theory*, Hungary (1975).
28. E. Pfaffelhuber, Learning and Information Theory, in *Int. J. Neurosci.* **3**:83 (1972); and in *Physical Principles of Neuronal and Organismic Behavior*, M. Conrad and M. Magar, eds., Gordon and Breach (1973).
29. H. Jeffreys, *Theory of Probability*, Oxford (1948).
30. W. T. Grandy, Fundamentals of Statistical Mechanics, preprint, University of Wyoming (1970).
31. H. B. Callen, *Thermodynamics*, Wiley (1960).
32. R. Kubo, K. Matsuo, and K. Kitahara, *J. Stat. Phys.* **9**:51 (1973).
33. M. Suzuki, *Prog. Theor. Phys.* **53**:1657 (1975).
34. J. Meixner, *Z. Phys.* **193**:366 (1966).
35. J. Meixner, *Rheol. Acta* **7**:8 (1968).
36. G. R. Gavalas, *Nonlinear Differential Equations of Chemically Reacting Systems*, Springer (1968).
37. G. R. Sell, *Topological Dynamics and Ordinary Differential Equations*, Van Nostrand (1971).