Statistical approach to the geometric structure of thermodynamics

Ryszard Mrugala*
Department of Mathematical Sciences, San Diego State University, San Diego, California 92182

James D. Nulton
Department of Mathematics, San Diego City College, San Diego, California 92101

J. Christian Schön and Peter Salamon
Department of Mathematical Sciences, San Diego State University, San Diego, California 92182
(Received 23 October 1989)

We show how both the contact structure and the metric structure of the thermodynamic phase space arise in a natural way from a generalized canonical probability distribution ρ. In particular, the metric form and the contact form are found to be derived from the microscopic entropy $s=-\ln\rho$. Thus the first law and the second law of thermodynamics can be given the geometric interpretation that a thermodynamic system must possess both a contact and a compatible metric structure. We proceed to construct explicitly a new nondegenerate bilinear form on the thermodynamic phase space, whose restriction to state space yields the Weinhold-Ruppeiner metric, and whose restriction to Gibbs space can serve as an alternative to the metric proposed by Gilmore.

I. INTRODUCTION

In 1973 Hermann[1] suggested that classical thermodynamics has its proper setting within the geometric framework of a thermodynamic phase space endowed with a contact structure. About the same time a second geometrical structure in the form of a Riemannian metric on the thermodynamic state space was introduced into thermodynamic theory by Weinhold.[2] In 1979 Ruppeiner[3] was able to use a related metric as the starting point for a new thermodynamic fluctuation theory. In 1983 Salamon and Berry[4] showed the intimate connection between this geometry and dissipation. Subsequent work[5–10] has elaborated our knowledge of both these aspects of geometry.

This paper sets forth a framework in which both of the above geometrical structures appear as natural statistical objects associated with a generalized canonical probability distribution ρ. In particular the contact form and the metric are shown to be the mean and the variance, respectively, of ds, where $s=-\ln\rho$ has the interpretation of microscopic entropy.

We offer this framework as a fertile formal setting for discussing the relation between geometry and thermodynamics. In this spirit we construct a nondegenerate bilinear form on the thermodynamic phase space whose restriction to state space gives the Ruppeiner metric, and whose construction is natural in that it is compatible with the statistical framework out of which both contact structure and Riemannian structure appear.

Furthermore, our bilinear form, when restricted to Gibbs space, provides a natural alternative to the metric constructed by Gilmore[7] in 1984.

II. STATISTICAL PRECURSORS OF GEOMETRY

A. The information-theoretic formulation of Jaynes

Let $\rho:\Gamma\rightarrow R^{+}$ be a non-normalized probability distribution on a space Γ. Any given functions $F^i:\Gamma\rightarrow R, i=1,\ldots,n$, may then be regarded as stochastic variables with respect to ρ.

Consider the problem of finding the function ρ that has maximum information-theoretic entropy $-\int\rho\ln\rho\, d\Gamma$ subject to the constraints

$$1=\int\rho\, d\Gamma\ , \tag{1}$$

$$x^i=\frac{\int F^i\rho\, d\Gamma}{\int\rho\, d\Gamma}\ . \tag{2}$$

This may be recognized as the maximum entropy formalism used by Jaynes[11] in his information-theoretic formulation of statistical mechanics.

The solution is well known and can be expressed in terms of the functions $F^1,\ldots,F^n$, whose mean values are constrained according to Eq. (2), and the Lagrange multipliers w and $p_1,\ldots,p_n$ corresponding to constraints expressed in (1) and (2), respectively,

$$\rho(\Gamma;w,p_1,\ldots,p_n)=\exp[-w+p_iF^i(\Gamma)]\ . \tag{3}$$

We could now proceed to exploit the normalization of ρ by expressing w as a function of the other Lagrange multipliers

$$w(p_1,\ldots,p_n)=\ln\int\exp(p_iF^i)d\Gamma\ , \tag{4}$$

whereupon the constraint Eqs. (2) become

$$x^i=\frac{\partial w}{\partial p_i}, \quad i=1,\ldots,n \ . \tag{5}$$

Indeed, these equations may be taken as the key to the physical interpretation of this formalism, since w can be interpreted as a thermodynamic potential with respect to which Eq. (4) is a fundamental relation and Eqs. (5) are the equations of state. We choose, however, to explicitly defer this normalization in the rest of this section, since our primary goal is to imbed our thermodynamic description in a space of larger dimension.

We take Eq. (3) as our starting point with ρ as a function of $n+1$ independent parameters and we define x^i from Eq. (2). That is,

$$x^i=\langle F^i\rangle=\frac{\int F^i\rho\, d\Gamma}{\int \rho\, d\Gamma} \ , \tag{6}$$

where at the same time we introduce $\langle \ \rangle$ to denote the expectation or mean over the space Γ.

By differentiating Eq. (6) we obtain

$$\frac{\partial x^i}{\partial p_j}=\langle F^iF^j\rangle - x^i x^j=\langle (F^i-x^i)(F^j-x^j)\rangle \ . \tag{7}$$

This equation offers a phenomenological interpretation of the covariances of the variables F^i. Alternatively, it gives a statistical interpretation of the derivatives $\partial x^i/\partial p_j$. Note that although ρ still has the free parameter w, the functions x^i in Eq. (6) depend only on $p_1,\ldots,p_n$ since the factor e^w appears in the numerator and the denominator of the expression on the right. Consequently, Eq. (7) is equivalent to

$$dx^i=\langle (F^i-x^i)(F^j-x^j)\rangle dp_j \ . \tag{8}$$

B. Microscopic entropy

The equations below define the microscopic entropy or bit number[6] of the distribution ρ and give its differential.

$$s=-\ln\rho = w - p_i F^i \ , \tag{9}$$

$$ds = dw - F^i dp_i \ . \tag{10}$$

These are both functions of the microscopic variables Γ [by way of $F^i(\Gamma)$], and of the parameters w and $p_1,\ldots,p_n$. Nevertheless, differentiation is understood to be only with respect to the variables $w,p_1,\ldots,p_n$.

We have called s the microscopic entropy, because its mean value is the macroscopic entropy. Schlögl calls s the bit number, and has connected the behavior of the cumulants of s with the onset of internal correlations[12] of the system. The quantity ds is closely related to what Schlögl calls the relative bit number. The latter's link with the Ruppeiner metric was explained in Ref. 6.

The mean and variance of ds lead naturally to the principal geometric constructions in this paper. In view of Eq. (6), the mean may be written

$$\langle ds\rangle = dw - x^i dp_i \ . \tag{11}$$

The variance $\langle (ds-\langle ds\rangle)^2\rangle$ is easily calculated using Eqs. (10) and (11) to be

$$\langle (F^i-x^i)(F^j-x^j)\rangle dp_i dp_j \ .$$

Equation (8) permits the more symmetrical form

$$\langle (ds-\langle ds\rangle)^2\rangle = dx^i dp_i \ . \tag{12}$$

Equation (11) leads to the contact structure of Sec. III, and Eq. (12) to the metric of Sec. IV.

III. CONTACT STRUCTURE

Consider now a $(2n+1)$-dimensional thermodynamic phase space M^{2n+1} whose independent coordinates are w, $x^1,\ldots,x^n$, and $p_1,\ldots,p_n$. In Eq. (11) the x^i are functions of $p_1,\ldots,p_n$, but we now regard them as independent, so that the form

$$\theta = dw - x^i dp_i \tag{13}$$

becomes a nondegenerate 1-form on M^{2n+1}, i.e., satisfies the condition

$$\theta\wedge(d\theta)^n\neq 0 \ . \tag{14}$$

This condition is just a coordinate-free statement of the requirement that on M^{2n+1} all coordinates w, $x^1,\ldots,x^n$, and $p_1,\ldots,p_n$ be independent. Such a pair (M^{2n+1},θ) where M^{2n+1} is a manifold and θ is a 1-form which fulfills Eq. (14) is called a contact manifold.

Of course, for real thermodynamic systems w, $x^1,\ldots,x^n$, and $p_1,\ldots,p_n$ are not independent. Indeed, the physical interpretation of these parameters requires the normalization of ρ, which, as noted in Sec. II, leads to Eqs. (4) and (5). In this context these $n+1$ conditions define an n-dimensional submanifold M^n of M^{2n+1}, on which the following equation holds:

$$\theta = dw - x^i dp_i = 0 \ . \tag{15}$$

Submanifolds of dimension n which satisfy Eq. (15) are called Legendre submanifolds of M^{2n+1}. In Sec. V we will give a general local characterization of these submanifolds.

Equation (15) can be seen as the complete Legendre transform of the Gibbs equation. This equation incorporates the first law of thermodynamics and includes the entropy as a function of state. However, as it places no convexity condition on the entropy, it does not incorporate the second law. Of course, the manifold M^n satisfies the second law by virtue of the maximum entropy principle from which it was derived. Nevertheless, we emphasize that this property does not come from the contact structure, rather it is related to the metric structure of Sec. IV.

The large space M^{2n+1} together with the contact structure given by θ offers the most natural setting for investigating the symmetry properties of thermodynamic systems. In Sec. V we will consider the Legendre transformations of thermodynamic coordinates. These transformations form a subgroup of the group of contact transformations, which are the most general transformations that preserve the 1-form θ and, thus, the Legendre submanifolds. Results of Sec. V also show that M^{2n+1}

offers the best way to describe a thermodynamic system without ascribing special importance to any particular choice of parameters.

IV. METRIC STRUCTURE

In this section we endow the contact space M^{2n+1} with yet another geometrical structure. The new structure is based on the variance of ds as given by Eq. (12). We again take x^i and p_i as independent variables. In this way $dx^i dp_i$ becomes a symmetric bilinear form on M^{2n+1}. This form is, however, degnerate; its rank is $2n$. We remove this degeneracy by adding the square of the contact form θ. This yields the nondegenerate symmetric bilinear form

$$G = dx^i dp_i + \theta^2 = dw\, dw - 2x^i dw\, dp_i + dx^j dp_j + x^i x^j dp_i dp_j \ . \quad (16)$$

G has the interpretation of a pseudo-Riemannian metric on M^{2n+1}. We observe that in the setting of Sec. II B Eqs. (11) and (12) show that G is the second moment $\langle (ds)^2 \rangle$ of ds.

Consider this metric restricted to the Legendre submanifold M^n described in Sec. III. On this submanifold an additional $n+1$ conditions hold between our variables

$$w = \Omega(p_1, \ldots, p_n) \ , \quad (17)$$

$$x^i = \frac{\partial \Omega}{\partial p_i}, \quad i = 1, \ldots, n \ . \quad (18)$$

When our metric G is restricted to this submanifold, we have

$$G|_{M^n} = g = \frac{\partial^2 \Omega}{\partial p_i \partial p_j} dp_i dp_j = dx^j dp_j \ . \quad (19)$$

Thus, on the submanifold M^n, the metric is equivalent to the Ruppeiner metric.[3] It is here that the connection may be seen between the metric and the second law of thermodynamics.[2] The function $\theta(p_1, \ldots, p_n)$ given by the right-hand side of Eq. (4) is easily seen to be convex. In fact, this may be taken as a statement of the second law. However, this convexity condition may be formulated geometrically as the requirement that $G|_S$ be positive definite for any Legendre submanifold S representing a physical system.

If we retain w as a free parameter and use only Eq. (18) to define a submanifold, we obtain an $(n+1)$-dimensional submanifold M^{n+1}. If G is restricted to M^{n+1}, we obtain

$$\begin{aligned} G|_{M^{n+1}} = \hat{g} &= dw\, dw - 2x^i dw\, dp_i + \frac{\partial x^j}{\partial p_i} dp_i dp_j \\ &\quad + x^i x^j dp_i dp_j \\ &= dw\, dw - 2\frac{\partial \Omega}{\partial p_i} dw\, dp_i \\ &\quad + \left[\frac{\partial^2 \Omega}{\partial p_i \partial p_j} + \frac{\partial \Omega \partial \Omega}{\partial p_i \partial p_j} \right] dp_i dp_j \ . \end{aligned} \quad (20)$$

The space M^{n+1} is an intensive counterpart of the Gibbs space. Therefore a complete Legendre transformation of $\hat{g}$ (see Sec. V) can be viewed as an alternative to the metric proposed by Gilmore.[7]

V. LEGENDRE TRANSFORMATIONS OF THE METRICS

All our constructions and discussions in the preceding sections were connected with the particular role ascribed to the parameters w and $p_1, \ldots, p_n$ and to the submanifolds M^n and M^{n+1} defined by Eqs. (5) and (6). We now consider the effect on our metric of allowing another choice of n independent variables from among the set $p_1, \ldots, p_n, x^1, \ldots, x^n$.

Formally, this can be achieved by considering the Legendre transformations of M^{2n+1}. These are a discrete subgroup of all transformations which leave the contact form θ invariant. A partial Legendre transformation on M^{2n+1} is given by the following $2n+1$ equations:

$$\begin{aligned} &W = w - p_I x^I, \quad P_I = -x^I, \quad P_J = p_J \ , \\ &X^I = p_I, \quad X^J = x^J \ . \end{aligned} \quad (21)$$

Here $I \cup J$ is a disjoint decomposition of the set of indices $\{1, \ldots, n\}$. θ retains its canonical appearance in these new coordinates,

$$\theta = dw - x^i dp_i = dW - X^i dP_i, \quad i = 1, \ldots, n \ . \quad (22)$$

To retain the same symbols for the same physical quantities, we now rewrite the terms $-X^i dP_i$ using our original parameters

$$\theta = dW + p_I dx^I - x^J dp_J \ . \quad (23)$$

We keep W because it can be interpreted as a partial Legendre transform of our original potential w.

The same procedure of first applying Eqs. (21) to G and then rewriting using our original variables gives us

$$G = dp_J dx^J - dx^I dp_I + (dW + p_I dx^I - x^J dp_J)^2 \ . \quad (24)$$

So far, the $2n+1$ variables x^I, p_I, x^J, p_J and W are all independent. As we have remarked earlier, Legendre submanifolds are spaces of equilibrium states of thermodynamic systems. The following useful theorem giving a local description of Legendre submanifolds is cited by Arnold.[13]

Theorem. For any partition $I \cup J$ of the set of indices $\{1, \ldots, n\}$ into two disjoint subsets and for a function $\Phi(x^I, p_J)$ of n variables x^i, $i \in I$, p_j, and $j \in J$, the $2n+1$ equations

$$p_I = -\frac{\partial \Phi}{\partial x^I} \ , \quad x^J = \frac{\partial \Phi}{\partial p_J} \ , \quad W = \Phi - x^I \frac{\partial \Phi}{\partial x^I} \quad (25)$$

define a Legendre submanifold S of M^{2n+1}. Conversely, every Legendre submanifold of M^{2n+1} is defined in a neighborhood of every point by these equations for at least one of the 2^n possible choices of the subset I.

We will use this result as motivation for a two-stage reduction of the metric G. The first reduction is accomplished by taking the variables p_I and x^J to be no longer independent but given as first derivatives of a potential Φ as in Eq. (25), while W is retained as an independent variable. This $n+1$ dimensional space can be thought of as a Legendre transform of Gibbs space. As a result of this, the restriction of G takes the form

$$\hat{g}_\Phi = -\frac{\partial^2\Phi}{\partial x^{I'}\partial x^I}dx^{I'}dx^I + \frac{\partial^2\Phi}{\partial p_J \partial p_{J'}}dp_J dp_{J'} + \left[dW - \frac{\partial\Phi}{\partial x^I}dx^I - \frac{\partial\Phi}{\partial p_J}dp_J\right]^2 . \tag{26}$$

For the second reduction, we take $W=\Phi(x^I,p_J)$ in accordance with Eqs. (21) and (25). With this step, the restriction of $\hat{g}_\Phi$ to n-dimensional Legendre submanifolds of θ becomes

$$g_\Phi = -\frac{\partial^2\Phi}{\partial x^{I'}\partial x^I}dx^{I'}dx^I + \frac{\partial^2\Phi}{\partial p_J \partial p_{J'}}dp_J dp_{J'} . \tag{27}$$

Observe that there are no terms of the form $dx\,dp$. This means that the nonzero terms in the metric matrix reside in diagonal blocks.

VI. CONCLUSIONS

The approach to thermodynamics based on contact geometry is a reformulation in which the elements of the classical theory assume their proper mathematical role. The Gibbs equation, which serves as a physical law applying to all systems, corresponds to the vanishing of the contact form θ. A fundamental relation and its attendant equations of state express the phenomenology of a particular system corresponding to a particular Legendre submanifold of thermodynamic phase space. This is analogous to phase space in classical dynamics, which supports a symplectic structure, and where a choice of Hamiltonian establishes the phenomenology.

A new formalism, if it is to be productive, should point beyond the theory it reformulates. The efforts of this paper were guided by the hope that clues to additional geometric structures important for macroscopic theory could be found by clarifying the microscopic origin of the contact form θ. Indeed, this form reveals itself as the mean or expected value of a quantity ds. The higher moments of this quantity should, therefore, also be important. In fact, the variance of this quantity is a bilinear form whose restriction to the Legendre submanifold of a physical system is a Riemannian metric whose significance in thermodynamic theory has already been recognized.[2-5,8-10] Moments beyond the second are related to the curvature associated with this Riemannian metric.

There is a compelling conceptual simplicity to this viewpoint. The particular metric associated with a given physical system is seen to be the restriction of the general pseudometric G. Like the contact form itself, G depends not on the particular system, but only on the choice of macroscopic parameters. It therefore participates, along with θ, in the embodiment of a physical law. Moreover, since the new structure G is intimately linked on the microscopic level with the contact structure, its introduction can be seen as a natural development in the same spirit as contact geometry.

Another theoretical bonus is derived from this point of view. The Gibbs equation, and therefore the contact geometry, does not address the second law of thermodynamics. As a result, some Legendre submanifolds cannot represent physical systems. The ones that are consistent with the second law may be characterized as those on which the restriction of the pseudometric G is a genuine semimetric, i.e., is positive semidefinite. Thus, the second law takes the form of a requirement of compatibility between the contact and the metric structures.

Three problems in connection with this theory require further study. One problem is to identify and to characterize the appropriate group of symmetries for the theory. One candidate is the subgroup of all contact transformations which preserve the metric.[14] Another is the group which preserves only the conformal structure associated with the metric, i.e., preserves the metric up to multiplication by a nonzero scalar function.

A second problem involves the interpretation of the metric G. The meaning of its restriction to a Legendre submanifold has been studied in earlier work,[2-5,8-10] but what is its role as a pseudometric in phase space? Can it provide a natural quantitative measure of the separation between Legendre submanifolds, and, therefore, between physical systems? We show elsewhere[15] that it is possible to obtain a Van der Waals-like real gas as the continuous deformation of an ideal gas by means of a one-parameter group of contact transformations. The "distance" between corresponding states of these two gases can be measured by applying G to the flows associated with this deformation. The result is a quantity which approaches zero asymptotically with large volume and low pressure, a rudimentary result which is at least consistent with any reasonable quantitative comparison of these two gases.

Related to this problem is the associated metric on M^{n+1} obtained in Sec. IV as a restriction of G. Since it reflects the properties of a system's potential surface, it can be expected to provide a theoretical tool for stability analysis of systems near equilibrium. We expect that the geometric apparatus of Secs. III and IV will be associated with any macroscopic theory derived on the basis of a maximum entropy principle as outlined in Sec. II A. Recently this maximum entropy formalism has been extended to the description of nonequilibrium phenomena.[16] Understanding the geometry in this context presents a third problem for future investigation.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Lars-Kai Hansen and Professor Stanislaw Sieniutycz for valuable discussions and comments. Dr. R. Mrugala and Dr. J. C. Schön would like to thank the Interdisciplinary Research Center for Scientific Modelling and Computation at San Diego State University for its hospitality. Finally, all authors gratefully acknowledge the fertile environment of the Telluride Summer Research Center where much of the final phase of the work took place.

*Permanent address: Institute of Physics, Nicolaus Copernicus University, PL-87-100 Torún, Poland.

[1]R. Hermann, *Geometry, Physics and Systems* (Dekker, New York, 1973).
[2]F. Weinhold, J. Chem. Phys. **63**, 2479 (1975); **63**, 2484 (1975); **63**, 2488 (1975); **63**, 2496 (1975); **65**, 559 (1976).
[3]G. Ruppeiner, Phys. Rev. A **20**, 1608 (1979).
[4]P. Salamon and S. Berry, Phys. Rev. Lett. **51**, 1127 (1983).
[5]G. Ruppeiner, Phys. Rev. A **27**, 1116 (1983).
[6]F. Schlögl, Z. Phys. B **59**, 449 (1985).
[7]R. Gilmore, Phys. Rev. A **30**, 1994 (1984).
[8]R. Mrugala, Rep. Math. Phys. **4**, 419 (1978).
[9]H. Janyszek and R. Mrugala, Rep. Math. Phys. **27**, 289 (1989).
[10]J. D. Nulton and P. Salamon, Phys. Rev. A **31**, 2520 (1985).
[11]E. T. Jaynes, Phys. Rev. **106**, 620 (1957).
[12]F. Schlögl, Z. Phys. B **58**, 57 (1985).
[13]V. I. Arnold, *Mathematical Methods of Classical Mechanics* (Springer-Verlag, New York, 1978), p. 367.
[14]P. Salamon, E. Ihrig, R. S. Berry, J. Math. Phys. **24** (10), 2515 (1983).
[15]R. Mrugala, J. D. Nulton, J. C. Schön, and P. Salamon, Rep. Math. Phys. (to be published).
[16]Walter T. Grandy, Jr., *Foundations of Statistical Mechanics* (Springer-Verlag, New York, 1988), Vol. II.