Riemannian structure of the thermodynamic phase space

G.F. Torres del Castillo
Departamento de Física Matemática, Instituto de Ciencias
Universidad Autónoma de Puebla, 72000 Puebla, Pue., México

AND

M. Montesinos Velasquez
Facultad de Ciencias Físico Matemáticas
Universidad Autónoma de Puebla
Apartado postal 1152, Puebla, Pue., México

Recibido el 8 de septiembre de 1992; aceptado el 14 de enero de 1993

ABSTRACT. Following Weinhold’s work, it is shown that it is possible to define a Riemannian metric on certain submanifolds of the space of equilibrium states of a thermodynamic system and that Weinhold’s abstract vector space can be identified with the tangent space to one of these submanifolds. It is also shown that the metric tensor can be written in terms of second derivatives of the internal energy, of the entropy or of other thermodynamic potentials.

RESUMEN. Siguiendo el trabajo de Weinhold, se muestra que es posible definir una métrica riemanniana en ciertas subvariedades del espacio de estados de equilibrio de un sistema termodinámico y que el espacio vectorial abstracto de Weinhold puede ser identificado con el espacio tangente a una de estas subvariedades. Se muestra también que el tensor métrico puede escribirse en términos de segundas derivadas de la energía interna, de la entropía o de otros potenciales termodinámicos.

PACS: 02.40.Ky; 05.70.-a

1. INTRODUCTION

The fact that each equilibrium state of a thermodynamic system can be characterized by means of a small number $n$ of independent parameters implies that the set of equilibrium states, hereafter called thermodynamic phase space, can be represented by points of $\mathbb{R}^n$. Even though the independent variables that are employed as coordinates of the thermodynamic phase space are usually restricted to be either extensive or intensive and to have some physical significance, there is a great arbitrariness in their choice and, therefore, the geometrical concepts, such as distances and angles, given by the representation of the thermodynamic phase space in $\mathbb{R}^n$ have no intrinsic meaning since they depend on the coordinates chosen. However, Weinhold [1-3] found that it is possible to define an intrinsic metric structure on a certain vector space associated with each equilibrium state of a thermodynamic system (see also Refs. [4-7]).

In this paper we show that a riemannian metric can be defined on certain submanifolds of the thermodynamic phase space whose dimensionality is given by the number of
independent intensive variables. Weinhold’s abstract space is isomorphic to the tangent space to one of these submanifolds at some equilibrium state. In Sect. 2 we summarize Weinhold’s construction of a vector space with scalar product associated with the thermodynamics of a given equilibrium state. In Sect. 3 we start from the law of increase of entropy to show that one can define a positive semidefinite, symmetric, second-rank tensor field on the thermodynamic phase space which induces riemannian metrics on certain submanifolds of that space (see also Ref. [7]) and in Sect. 4 we give some examples of these metrics.

2. SUMMARY OF WEINHOLD’S CONSTRUCTION

Assuming that the internal energy $U$ is expressed as a differentiable function of $r$ extensive state functions $X^1, X^2, \ldots, X^r$ where $r$ is fixed by the Gibbs phase rule, the field variables $R_i$, conjugate to $X^i$, are defined by

$$R_i \equiv \frac{\partial U}{\partial X^i}. \quad (1)$$

Then, with each field differential $dR_i$ Weinhold associates an “abstract vector” $|R_i\rangle$ and defines a scalar product among these vectors through

$$\langle R_i | R_j \rangle \equiv \left. \frac{\partial R_i}{\partial X^j} \right|_\xi, \quad (2)$$

where the subscript $\xi$ denotes that the partial derivative is to be evaluated at a particular state of interest. By expressing the second law of thermodynamics through the condition

$$\frac{\partial^2 U}{\partial (X^i)^2} \bigg|_\xi = \left. \frac{\partial R_i}{\partial X^j} \right|_\xi \geq 0, \quad (i \text{ not summed}) \quad (3)$$

from Eq. (2) it follows that $\langle R_i | R_i \rangle \geq 0$ and, assuming that the field variables $R_i$ are independent, Weinhold concludes that $\langle R_i | R_i \rangle = 0$ only if $|R_i\rangle = 0$. According to Eqs. (1-2), the scalar product $\langle R_i | R_j \rangle$ can be expressed as

$$\langle R_i | R_j \rangle = \left. \frac{\partial^2 U}{\partial X^j \partial X^i} \right|_\xi, \quad (4)$$

which shows that the scalar product (2) satisfies $\langle R_i | R_j \rangle = \langle R_j | R_i \rangle$. It may be noticed that Eq. (2) or (4) defines only the scalar product among the vectors $|R_i\rangle$, which can be regarded as basis vectors of a certain vector space; the symmetry and the bilinearity of this scalar product have to be imposed additionally.

As we shall show in the next section, Weinhold’s abstract vector space can be identified with the tangent space to a submanifold of the thermodynamic phase space at a particular state.
3. RIEMANNIAN STRUCTURES

We shall assume that the entropy of a given thermodynamic system can be expressed in terms of \( n \) extensive independent state variables \( Y^1, \ldots, Y^n \), \( S = S(Y^1, \ldots, Y^n) \), and that \( S \) is a differentiable function of the \( Y^i \); then, using the fact that for an isolated system the entropy does not decrease when any constraint is removed, it follows that

\[
\frac{\partial^2 S}{\partial Y^i \partial Y^j} a^i a^j \leq 0
\]  

(5)

for all \( a^i \), where, as in what follows, there is summation over repeated indices. Indeed, let us consider an arbitrary equilibrium state for which the variables \( Y^i \) take the values \( Y^i_0 \) and let us assume that the system is divided, by means of appropriate walls, into two subsystems characterized by the values of the state variables \( \frac{1}{2} Y^i_0 + \lambda a^i \) and \( \frac{1}{2} Y^i_0 - \lambda a^i \), respectively, where \( \lambda \) is a small parameter, in such a way that, owing to the extensive character of the variables \( Y^i \), the total value of \( Y^i \) for the composite system is \( Y^i_0 \). The entropy of each subsystem is, to second order in \( \lambda \),

\[
S \left( \frac{1}{2} Y^1_0, \ldots, \frac{1}{2} Y^n_0 \right) \pm \lambda a^i \frac{\partial S}{\partial Y^i} \bigg|_{(\frac{1}{2} Y^i_0)} + \frac{1}{2} \lambda^2 a^i a^j \frac{\partial^2 S}{\partial Y^i \partial Y^j} \bigg|_{(\frac{1}{2} Y^i_0)} + \cdots.
\]

Therefore, using the fact that \( S \left( \frac{1}{2} Y^1_0, \ldots, \frac{1}{2} Y^n_0 \right) = \frac{1}{2} S \left( Y^1_0, \ldots, Y^n_0 \right) \) and also that \( (\partial^2 S/\partial Y^i \partial Y^j) \bigg|_{(\frac{1}{2} Y^i_0)} = 2(\partial^2 S/\partial Y^i \partial Y^j) \bigg|_{(Y^i_0)} \), we find that the entropy of the composite system is

\[
S(Y^1_0, \ldots, Y^n_0) + 2\lambda^2 a^i a^j \frac{\partial^2 S}{\partial Y^i \partial Y^j} \bigg|_{(Y^i_0)} + \cdots.
\]  

(6)

Thus, if the composite system is isolated and the constraints are removed, the entropy of the system will become \( S(Y^1_0, \ldots, Y^n_0) \), which must be greater than, or equal to, the total entropy (6); from which the inequality (5) follows.

The functions \( \partial^2 S/\partial Y^i \partial Y^j \) can be regarded as the components of a symmetric, negative semidefinite, second-rank tensor field:

\[
h = \frac{\partial^2 S}{\partial Y^i \partial Y^j} dY^i dY^j = d \left( \frac{\partial S}{\partial Y^i} \right) dY^i,
\]

(7)

where, as in the forthcoming, juxtaposition of differentials means symmetrized tensor product. On the other hand, \( dS \) can be expressed in the form

\[
dS = \frac{1}{T} dU - \frac{1}{T} \sum_{i=2}^n F_i dX^i,
\]

(8)
where $T$ is the absolute temperature, $U$ is the internal energy and the $X^i$ ($i = 2, \ldots, n$) are extensive variables. Choosing $Y^1 \equiv U$ and $Y^i \equiv X^i$ for $i = 2, \ldots, n$, from Eqs. (7–8) we find that $\partial S/\partial Y^1 = 1/T$ and $\partial S/\partial Y^i = -F_i/T$ for $i = 2, \ldots, n$; thus

$$h = d\left(\frac{1}{T}\right) dU + \sum_{i=2}^{n} d\left(-\frac{F_i}{T}\right) dX^i$$

$$= -\frac{1}{T^2} dT \left( dU - \sum_{i=2}^{n} F_i dX^i \right) - \frac{1}{T} \sum_{i=2}^{n} dF_i dX^i$$

$$= -\frac{1}{T} \left( dT dS + \sum_{i=2}^{n} dF_i dX^i \right)$$

Hence, assuming $T > 0$, we conclude that

$$g \equiv -Th = dT dS + \sum_{i=2}^{n} dF_i dX^i$$

is a symmetric, positive semidefinite, second-rank tensor field (see also Ref. [8]).

Equation (8) gives $dU = T dS + \sum_{i=2}^{n} F_i dX^i$; therefore, regarding $U$ as a function of the extensive variables $X^1 = S, X^2, \ldots, X^n$, it follows that $\partial U/\partial X^1 = T$ and $\partial U/\partial X^i = F_i$ for $i = 2, \ldots, n$. Thus, from Eq. (9) we obtain

$$g = d\left(\frac{\partial U}{\partial X^1}\right) dX^1 + \sum_{i=2}^{n} d\left(\frac{\partial U}{\partial X^i}\right) dX^i = \frac{\partial^2 U}{\partial X^i \partial X^j} dX^i dX^j.$$

The positive semidefiniteness of $g$ amounts to the condition

$$\frac{\partial^2 U}{\partial X^i \partial X^j} a^i a^j \geq 0,$$

for all $a^i$. This last condition is usually taken as the starting point in the definition of a metric in the thermodynamic phase space (cf. Eq. (3)). The equality in Eq. (11) does not imply $a^i = 0$ (compare Refs. [5,6]); in fact, the homogeneity of $U(X^1, \ldots, X^n)$ implies the Gibbs-Duhem relation $0 = X^i d(\partial U/\partial X^i) = X^i (\partial^2 U/\partial X^i \partial X^j) dX^j$, which, owing to the linear independence of the $dX^j$, yields

$$X^i \frac{\partial^2 U}{\partial X^i \partial X^j} = 0,$$

therefore,

$$\det \left(\frac{\partial^2 U}{\partial X^i \partial X^j}\right) = 0$$

(see also Ref. [7]).
In summary, if \( dU = T \, dS + \sum_{i=2}^{n} F_i \, dX^i \), then the symmetric, second-rank tensor field \( g = dT \, dS + \sum_{i=2}^{n} dF_i \, dX^i \) is positive semidefinite (degenerate). Nevertheless, the restriction of \( g \) to certain submanifolds of the thermodynamic phase space is positive definite (see also Ref. [7]).

Following Ref. [1] we denote

\[
R_i \equiv \frac{\partial U}{\partial X^i}, \quad (i = 1, \ldots, n)
\]

which are intensive variables. Then, Eq. (13) amounts to \( \det(\partial R_i / \partial X^j) = 0 \), which means that the \( n \) intensive variables \( R_i \) are dependent (this also follows from the Gibbs-Duhem relation). Let \( r \) be the rank of the matrix \( (\partial R_i / \partial X^j) \); in other words, only \( r \) of the variables \( R_i \) are independent \( (r < n) \). By renaming the variables if necessary, we can assume that \( R_1, \ldots, R_r \) are independent; therefore, if the Greek indices \( \alpha, \beta, \ldots \), range and sum over \( 1, \ldots, r \), \( \det(\partial R_\alpha / \partial X^\beta) = \det(\partial^2 U / \partial X^\alpha \partial X^\beta) \neq 0 \) (this follows from the symmetry of \( \partial R_i / \partial X^j \)) and

\[
dl^2 = g_{\alpha \beta} dX^\alpha \, dX^\beta \equiv \frac{\partial^2 U}{\partial X^\alpha \partial X^\beta} \, dX^\alpha \, dX^\beta = \frac{\partial R_\alpha}{\partial X^\beta} \, dX^\alpha \, dX^\beta = dR_\alpha \, dX^\alpha,
\]

is a riemannian metric (i.e., a symmetric, positive definite, second-rank tensor field) on each submanifold of dimension \( r \) defined by \( X^{r+1} = \text{const.}, \ldots, X^n = \text{const.} \) or, equivalently, by \( dX^{r+1} = \cdots = dX^n = 0 \). Comparison with Eq. (10) shows that \( dl^2 \) is the restriction of \( g \) to a submanifold \( dX^{r+1} = \cdots = dX^n = 0 \); therefore, from Eq. (9) we get

\[
dl^2 = dT \, dS + \sum_{i=2}^{r} dF_i \, dX^i.
\]

It may be pointed out that \( r \) can be less than \( n - 1 \) (compare Ref. [7]). Equivalently, the number of linearly independent null vectors with respect to \( g \) (i.e., the vectors \( a^i \) satisfying \( a^i(\partial^2 U / \partial X^i \partial X^j) = 0 \)) may be greater than 1. The tensor field \( g \) is positive definite when restricted to any submanifold transversal to the null vectors of \( g \), which needs not be given by equations of the form \( dX^1 = 0 \).

In any riemannian manifold the gradient of a scalar function \( f \) defined on it can be defined as the vector field with components

\[
(\text{grad} f)^\beta = g^{\alpha \beta} \frac{\partial f}{\partial X^\alpha},
\]

where \( (g^{\alpha \beta}) \) is the inverse of the matrix \( (g_{\alpha \beta}) \) formed with the components of the metric tensor with respect to a coordinate system \( X^1, X^2, \ldots \). In the present case, each intensive variable \( R_\alpha \) restricted to a submanifold \( X^{r+1} = \text{const.}, \ldots, X^n = \text{const.} \), is a function of the \( r \) variables \( X^\beta \) and, according to Eqs. (15) and (17), the components of the gradient of \( R_\alpha \) are given by

\[
(\text{grad} \, R_\alpha)^\beta = g^{\beta \gamma} \frac{\partial R_\alpha}{\partial X^\gamma} = g^{\beta \gamma} g_{\alpha \gamma} = \delta_\alpha^\beta,
\]

(18)
therefore the scalar product of the vector fields $\text{grad } R_\alpha$ and $\text{grad } R_\beta$ is

$$\text{grad } R_\alpha \cdot \text{grad } R_\beta = g_{\gamma\epsilon}(\text{grad } R_\alpha)\gamma(\text{grad } R_\beta)\epsilon$$

$$= g_{\gamma\epsilon}\delta^\gamma_\alpha\delta^\epsilon_\beta = g_{\alpha\beta} = \frac{\partial R_\alpha}{\partial X^\beta} = \frac{\partial^2 U}{\partial X^\alpha \partial X^\beta}. \quad (19)$$

Comparison with Eq. (4) shows that the abstract vector space considered by Weinhold [1-3], spanned by the vectors $|R_\alpha\rangle$, can be identified with the tangent space to the submanifold $X^{r+1} = \text{const.}, \ldots, X^n = \text{const.}$ at a particular state and that the abstract vector $|R_\alpha\rangle$ corresponds to the gradient of $R_\alpha$ evaluated at that particular state.

In a similar manner, the components of the gradient of $X^\alpha$ are

$$(\text{grad } X^\alpha)^\beta = g^{\beta\gamma}\frac{\partial X^\alpha}{\partial X^\gamma} = g^{\alpha\beta}, \quad (20)$$

hence,

$$\text{grad } R_\alpha \cdot \text{grad } X^\beta = g_{\gamma\epsilon}(\text{grad } R_\alpha)\gamma(\text{grad } X^\beta)\epsilon$$

$$= g_{\gamma\epsilon}\delta^\gamma_\alpha g^{\beta\epsilon} = \delta^\beta_\alpha, \quad (21)$$

and

$$\text{grad } X^\alpha \cdot \text{grad } X^\beta = g_{\gamma\epsilon}(\text{grad } X^\alpha)\gamma(\text{grad } X^\beta)\epsilon$$

$$= g_{\gamma\epsilon} g^{\alpha\gamma} g^{\beta\epsilon} = g^{\alpha\beta}. \quad (22)$$

Furthermore, since $g_{\alpha\beta} = \partial R_\alpha / \partial X^\beta$, by virtue of the chain rule,

$$g^{\alpha\beta} = \frac{\partial X^\alpha}{\partial R_\beta}, \quad (23)$$

where $X^\alpha$ is expressed as a function of the $R_\gamma$. Thus,

$$\text{grad } X^\alpha \cdot \text{grad } X^\beta = \frac{\partial X^\alpha}{\partial R_\beta} \quad (24)$$

[cf. Eq. (19)]. The symmetry of $g^{\alpha\beta}$ amounts to $\partial X^\alpha / \partial R_\beta = \partial X^\beta / \partial R_\alpha$, which implies that, locally, the variables $X^\alpha$ can be expressed in the form

$$X^\alpha = \frac{\partial \Phi}{\partial R_\alpha}, \quad (25)$$

where $\Phi$ is some function of the $R_\alpha$ (cf. Eq. (14)). From Eqs. (23) and (25) we obtain

$$g^{\alpha\beta} = \frac{\partial^2 \Phi}{\partial R_\alpha \partial R_\beta} \quad (26)$$
(cf. Eq. (15) and Ref. [9]) and Eq. (15) yields

\[ d\ell^2 = dR_\alpha d \left( \frac{\partial \Phi}{\partial R_\alpha} \right) = \frac{\partial^2 \Phi}{\partial R_\alpha \partial R_\beta} dR_\alpha dR_\beta. \]  

(27)

4. Examples

As shown in Refs. [3] and [5], the existence of a scalar product in the vector space associated with each equilibrium state allows one to obtain thermodynamic relations from the Cauchy-Schwarz and Bessel inequalities. As an example, we shall consider the case of a thermodynamic system for which \( dU = TdS - PdV + \mu dN \), where \( P, V, \mu, \) and \( N \) are the pressure, volume, chemical potential, and mole number. From Eq. (16) we see that

\[ d\ell^2 = dT dS - dP dV \]  

(28)

is a riemannian metric on each two-dimensional submanifold \( N = \text{const.} \), provided \( P \) and \( T \) are independent. By choosing the variables \( T \) and \( V \) as coordinates on a submanifold \( N = \text{const.} \) and using the Maxwell relation \( \left( \frac{\partial S}{\partial V} \right)_T, N = \left( \frac{\partial P}{\partial T} \right)_V, N \), one finds that Eq. (28) amounts to \( d\ell^2 = \left( \frac{\partial S}{\partial T} \right)_{V,N} (dT)^2 - \left( \frac{\partial P}{\partial V} \right)_{T,N} (dV)^2 \) (this means that the coordinate system \( (T, V) \) is orthogonal), equivalently,

\[ d\ell^2 = C_V \frac{(dT)^2}{T} + \frac{1}{\kappa_T V^2} (dV)^2, \]  

(29)

where \( C_V \) is the heat capacity at constant volume and \( \kappa_T \equiv -\frac{1}{V} \left( \frac{\partial V}{\partial P} \right)_T \) is the isothermal compressibility (cf. Ref. [10]). The positive definite character of the metric (29) is equivalent to the stability conditions \( C_V > 0, \kappa_T > 0 \) and to \( \kappa_T < \infty \). (The fact that \( \kappa_T \) may become infinite at some points does not contradict our conclusions since at those points \( P \) and \( T \) are not independent.)

If \( \{v_1, v_2\} \) is an orthogonal basis of a vector space and \( w \) is an arbitrary vector, Bessel’s inequality yields \( w \cdot w = (w \cdot v_1)^2 + (w \cdot v_2)^2 \). Therefore, using that \( \text{grad} \ T \) and \( \text{grad} \ V \) are orthogonal [see Eq. (21)] and applying the foregoing identity to \( \text{grad} \ S \), from Eqs. (19), (21) and (24) one finds

\[ \frac{C_P}{T} = \frac{C_V}{T} + \frac{\beta^2 V}{\kappa_T}, \]  

where \( C_P \) is the heat capacity at constant pressure and \( \beta \equiv \frac{1}{V} \left( \frac{\partial V}{\partial T} \right)_P \) is the thermal expansion coefficient. Similarly, taking \( v_1 = \text{grad} \, P, \ v_2 = \text{grad} \, S \) (which, in view of Eq. (21), are orthogonal) and \( w = \text{grad} \, V \) one obtains

\[ V \kappa_T = V \kappa_S + \frac{\beta^2 V^2 T}{C_P}, \]
where \( \kappa_S = -\frac{1}{V} \left( \frac{\partial V}{\partial P} \right)_s \) is the adiabatic compressibility. Additional relations are derived in Refs. [3,5,11]. Using Eq. (25) we find that, in the present case, \( d\Phi = X^\alpha dR_\alpha = SdT - VdP \); therefore, \( \Phi = -G \left|_{N=\text{const.}} \right. \), where \( G \) is the Gibbs function.

Alternatively, on each submanifold \( V = \text{const.} \),

\[
d\ell^2 = dT dS + d\mu dN
\]

is a riemannian metric, provided \( T \) and \( \mu \) are independent. Choosing \( T \) and \( N \) as coordinates and using the Maxwell relation \( \left( \frac{\partial S}{\partial N} \right)_{T,V} = -\left( \frac{\partial \mu}{\partial T} \right)_{N,V} \) one finds that Eq. (30) is equivalent to

\[
d\ell^2 = C_V \left( \frac{dT}{T} \right)^2 + \left( \frac{\partial \mu}{\partial N} \right)_{T,V} (dN)^2.
\]

Now, \( d\Phi = SdT + N d\mu \); hence, \( \Phi = -\Omega \left|_{V=\text{const.}} \right. \), where \( \Omega \) is the thermodynamic potential given by \( \Omega = -PV \).

The expression \( dT dS - dP dV \) appearing in the right-hand side of Eq. (28) corresponds to the maximum amount of useful work that can be extracted from a thermodynamic system immersed in a bath at temperature \( T + dT \) and subjected to an external pressure \( P + dP \), where \( T \) and \( P \) are the temperature and pressure of the system [12] (see also Ref. [4]). The right-hand side of Eq. (16) has a similar significance.

The existence of a riemannian metric on submanifolds of the thermodynamic phase space allows us to introduce various geometric notions such as length of curves, geodesics, parallel translation and curvature. For instance, in the case of a monatomic ideal gas, the metrics (29) and (31) become

\[
dl^2 = C_V \left( \frac{dT}{T} \right)^2 + \left( \frac{RT}{V^2} \right) (dV)^2, \quad (N = \text{const.})
\]

and

\[
dl^2 = C_V \left( \frac{dT}{T} \right)^2 + \left( \frac{RT}{N} \right) (dN)^2, \quad (V = \text{const.})
\]

(see also Ref. [7]). A straightforward computation shows that the metrics (32) are flat. On the other hand, for a van der Waals gas (assuming \( C_V \) constant and \( N = 1 \)), the metric (29) takes the form

\[
dl^2 = C_V \left( \frac{dT}{T} \right)^2 + \left[ \frac{RT}{(V - b)^2} - \frac{2a}{V^3} \right] (dV)^2.
\]

The nonvanishing components of the curvature corresponding to this metric are determined by the gaussian curvature

\[
K = \frac{RAV^3(V - b)^2}{2C_V[RTV^3 - 2a(V - b)^2]^2},
\]

which is positive.
5. CONCLUDING REMARKS

The existence of a metric associated with the thermodynamic phase space allows one to give a geometric interpretation to various thermodynamic relations that are usually obtained by other means. This metric also allows one to define the concept of length for fluctuations about equilibrium states (see, e.g., Refs. [5,13]). However, it is not clear to what extent it is possible to establish a correspondence between geometric and thermodynamic concepts. It may be pointed out that the thermodynamic relations derived in Refs. [3,5] and in the preceding section are conformally invariant in the sense that they are unchanged if the metric tensor $d\ell^2$ is replaced by $\phi^2 d\ell^2$, where $\phi$ is any nonvanishing real-valued function.

ACKNOWLEDGEMENTS

The authors are grateful to Dr. E. Cantoral Uriza and P. Martínez Garcilazo for useful comments and to the referee for bringing to their attention Refs. [4,6-8,13].

REFERENCES

9. H.B. Callen, Thermodynamics, Wiley, New York (1960); Appendix G.