Journal of Statistical Physics, Vol. 42, Nos. 5/6, 1986

# A Geometrical Measure for Entropy Changes

**Tova Feldmann,[1] R. D. Levine,[1] and Peter Salamon[2]**

*Received September 4, 1984; revision received September 6, 1985*

The geometrical approach to statistical mechanics is used to discuss changes in entropy upon sequential displacements of the state of the system. An interpretation of the angle between two states in terms of entropy differences is thereby provided. A particular result of note is that any state can be resolved into a state of maximal entropy (both states having the same expectation values for the constraints) and an orthogonal component. A cosine law for the general case is also derived.

**KEY WORDS**: Entropy changes; statistical mechanics.

## 1. INTRODUCTION

The geometrical approach to thermodynamics[1] has centered attention on equilibrium states. More recently, it has been applied to processes[2,3] and has been generalized to systems not in equilibrium.[4–8] Our intention here is to consider a general process through a sequence of arbitrary states and relate the fundamental new notion of the geometrical approach, namely, the angle between two states, to the change in entropy.

When we adopt a statistical description where the physical state of the system is given uniquely by specifying a probability distribution $\{p^i, i=1,\ldots, N\}$ over the $N$ possible, mutually exclusive, and collectively exhaustive states. In the geometrical approach the scalar product of two states is given by

$$\mathbf{p}\cdot\mathbf{q}=\sum_i\sum_j g_{ij}\,p^i q^j \tag{1}$$

[1] Department of Physical Chemistry, The Hebrew University, Jerusalem 91904, Israel.
[2] Department of Mathematical Sciences, San Diego State University, San Diego, California 92182.

where $g_{ij}$ is the metric tensor for the space of states.[4–8] It proves convenient to regard the space of states as Euclidean and this can always be arranged by using an $N$-dimensional space.[7] The physical states are then confined to an $(N-1)$-dimensional subspace (the unit hypersphere[7]) by the requirement of normalization

$$\mathbf{p} \cdot \mathbf{p} = 1 \tag{2}$$

On the subspace of physical states, the shortest distance[8] $d$ between two states $\mathbf{p}$ and $\mathbf{q}$ is the angle between them[7]:

$$\begin{aligned} d &= \cos^{-1}(\mathbf{p} \cdot \mathbf{q}) \\ &= \cos^{-1}\left[\sum_i (p^i)^{1/2}(q^i)^{1/2}\right] \end{aligned} \tag{3}$$

Locally, the metric is given by

$$g_{ij} = \delta_{ij}/p^j \tag{4}$$

It can be regarded as a reflection of the convexity property of the information theoretic entropy $S = -\sum_i p^i \ln p^i$, i.e.,

$$g_{ij} = -\partial^2 S/\partial p^i \, \partial p^j \tag{5}$$

Figure 1 which shows contours connecting states of equal entropy on the unit sphere in $N = 3$ dimensions is useful in elucidating this aspect. In Sections 2 and 3 we shall relate the changes in entropy to the geometry on the sphere. In particular we note from (4) and (1) that the scalar product of two independent infinitesimal displacements $\delta p^i$ and $\delta q^i$ both originating from the same state $p^{0i}$ is given by

$$\delta\mathbf{p} \cdot \delta\mathbf{q} = \sum_i \frac{\delta p^i \, \delta q^i}{p^{0i}} \tag{6}$$

## 2. ENTROPY

The quantity of interest in discussing availability and dissipation is the entropy deficiency[9,10] between the distributions $\mathbf{p}$ and $\mathbf{p}^0$:

$$DS(\mathbf{p}, \mathbf{p}^0) = \sum_i p^i \ln(p^i/p^{0i}) \tag{7}$$

In vector notation, the entropy is the scalar product of the state vector $\mathbf{p}$ and the surprisal, $I_i = -\ln p^i$, which is a vector in the dual space[7] of observables

$$S = \mathbf{I} \cdot \mathbf{p} \tag{8}$$

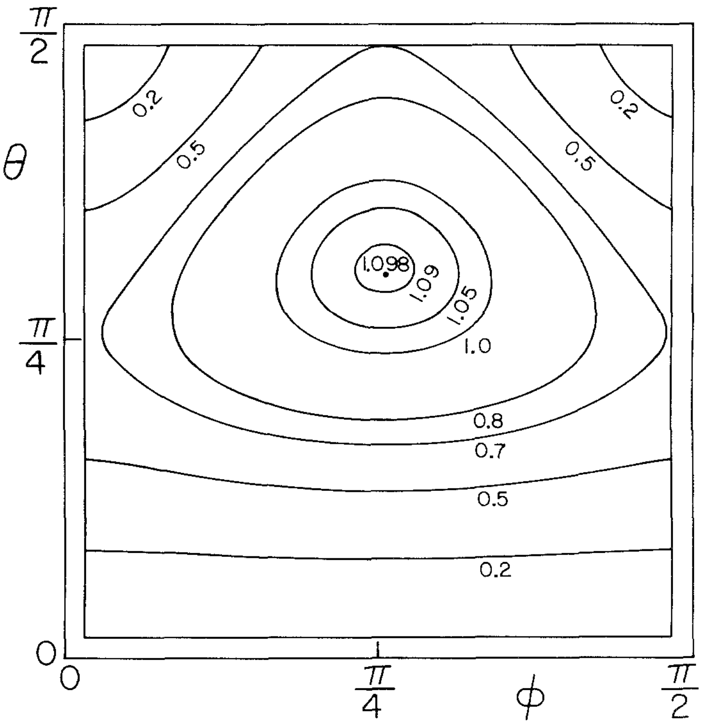


Fig. 1. Contour plot of the entropy for a normalized distribution of three states ($N = 3$). Shown are contours connecting states of equal entropy, with the numerical value of the entropy as given in the drawing. The proper plot should be over the first octant of the unit sphere in three-dimensional Cartesian space $p_i = x_i^2$, $S = -2\sum_i x_i^2 \ln x_i$, $i = 1, 2, 3$, $x_i \geqslant 0$. Shown instead is a projection defined in terms of $\phi$ and $\theta$ (the angular coordinates of a point on the unit sphere, $x_1 = \sin\theta\cos\phi$, $x_2 = \sin\theta\sin\phi$, $x_3 = \cos\theta$).

Similarly

$$DS = (\mathbf{I}_0 - \mathbf{I}) \cdot \mathbf{p} \tag{9}$$

Consider now three different displacements of state. One between $\mathbf{p}$ and $\mathbf{p}^0$, one between $\mathbf{q}$ and $\mathbf{p}^0$ and one between $\mathbf{p}$ and $\mathbf{q}$. Then

$$\Delta\, DS \equiv DS(\mathbf{q}, \mathbf{p}^0) - [DS(\mathbf{q}, \mathbf{p}) + DS(\mathbf{p}, \mathbf{p}^0)] \tag{10}$$

For small displacements where

$$\begin{aligned} \mathbf{p} &= \mathbf{p}^0 + \delta\mathbf{p} \\ \mathbf{q} &= \mathbf{p}^0 + \delta\mathbf{p} + \delta\mathbf{q} \end{aligned} \tag{11}$$

one obtains to second order in the displacements

$$\delta\, DS = \delta\mathbf{p} \cdot \delta\mathbf{q} \equiv \sum_i \frac{\delta p^i\, \delta q^i}{p^{0i}} \tag{12}$$

The result (12) is the relation between the change in entropy and the angle between two consecutive displacements of the state. In Section 3 we examine three particular cases: (1) orthogonal displacements $\delta\mathbf{p} \cdot \delta\mathbf{q} = 0$, (2) displacements along a geodesic, and (3) general consecutive steps where we interpret (12) as a "cosine law" on the unit (hyper)sphere.

## 3. DISPLACEMENTS

We consider the angles between different possible displacement of states.

### 3.1. Orthogonal Displacements

The first results is of interest in the variety of contexts wherein one represents the distribution of one of maximal entropy subject to one or more constraints.[11] We shall work with one constraint only since it simplifies the notation and the required generalization to many constraints is rather obvious.

Let $\mathbf{q}$ and $\mathbf{p}$ be two distinct normalized distributions which are both consistent with the same constraint, that is,

$$\mathbf{A} \cdot \mathbf{q} = \sum_i A_i q^i = \langle A \rangle \tag{13a}$$

$$\mathbf{A} \cdot \mathbf{p} = \sum_i A_i p^i = \langle A \rangle \tag{13b}$$

here $A_i$ is the value of the observable $A$ for the state $i$ and we emphasize that the numerical value, $\langle A \rangle$, of the expectation of the observable is the same in (13a) and (13b). To insure that $\mathbf{p}$ and $\mathbf{q}$ are distinct the reader may wish to take it[12] that there is at least one nontrivial observable $B$ such that

$$\mathbf{B} \cdot \mathbf{p} \neq \mathbf{B} \cdot \mathbf{q} \tag{14}$$

However, we shall not use (14) explicitly in what follows.

Our second condition is that both $\mathbf{p}$ and $\mathbf{p}^0$ are normalized distributions whose entropy is maximal, subject to a given expectation value (not quite the same) of $A$:

$$\begin{aligned} \langle A \rangle &= \mathbf{A} \cdot \mathbf{p} = \mathbf{A} \cdot \mathbf{p}^0 + \delta\langle A \rangle \\ \delta\langle A \rangle &= \sum_i A_i \, \delta p^i \end{aligned} \tag{15}$$

By our condition

$$\begin{aligned} p^i &= \exp[-(\lambda_0 + \delta\lambda_0) - (\lambda + \delta\lambda) A_i] \\ p^{0i} &= \exp[-\lambda_0 - \lambda A_i] \end{aligned} \tag{16}$$

Here $\lambda_0$ is the Lagrange multiplier for normalization. To compute $\delta p^i = p^i - p^{0i}$ we note that since $\sum \delta p^i = 0$ and $\delta\lambda$ is small it follows that

$$\delta\lambda_0 = -\langle A \rangle \, \delta\lambda \tag{17}$$

and so, from (16)

$$(\delta p^i / p^{0i}) = -\delta\lambda(A_i - \langle A \rangle) \approx \ln(p^i / p^{0i}) \tag{18}$$

Using (13), (18), and the definition (1) of the scalar product of two vectors we obtain from (10)

$$\begin{aligned} \Delta\, DS &\equiv [I(\mathbf{p}^0) - I(\mathbf{p})] \cdot (\mathbf{q} - \mathbf{p}) \\ &= \delta\mathbf{p} \cdot (\mathbf{q} - \mathbf{p}) = 0 \end{aligned} \tag{19}$$

The orthogonality of $\mathbf{q} - \mathbf{p}$ to $\delta\mathbf{p}$, as stated in (19) is a central result of this paper. It states the following: A distribution is a point in our space. Consider the set of normalized and unique distributions which are of maximal entropy subject to the expectation value of the observable $A$. As $\langle A \rangle$ varies, the distribution changes in a continuous fashion and hence by varying $\langle A \rangle$ we trace a continuous curve. Other distributions (such as $\mathbf{q}$) which have a given value of $\langle A \rangle$ are always orthogonal to a small displacement $\delta\mathbf{p}$ along the curve, $(\mathbf{q} - \mathbf{p}) \cdot \delta\mathbf{p} = 0$.

A special case of our result is when $\mathbf{p}$ is a distribution of maximal entropy subject to a given value, $\langle A \rangle$, of the expectation of $A$, and $\mathbf{q}$ is a distribution of maximal entropy subject to given values of $\langle A \rangle$ and of $\langle B \rangle$, $\langle B \rangle = \mathbf{B} \cdot \mathbf{q}$ such that (14) holds. For that case the result $\Delta\, DS = 0$ was derived by Hobsen and Cheng.[12] It leads to a special version of our conclusion: In the maximum entropy formalism the addition of a constraint always shifts the distribution in a perpendicular direction. This observation is relevant to the description of relaxation in terms of distributions of maximal entropy[13] with a diminishing number of constraints.[14]

### 3.2. Parallel Displacements

From orthogonal we turn to parallel displacements, that is, to the distance $d_{20}$ from $\mathbf{p}^0$ to $\mathbf{q}$ being the sum

$$d_{20} = d_{21} + d_{10} \tag{20}$$

of the distances from $\mathbf{p}^0$ to $\mathbf{p}$ and from $\mathbf{p}$ to $\mathbf{q}$, respectively. Since our normalized distributions are points on the unit (hyper) sphere[7] we expect that (20) will be satisfied when the three distributions lie along a geodesic. This is indeed so, almost by definition, and so we provide a proof only for the simplest case, that of three state. On the sphere, $\mathbf{p}^0$, $\mathbf{p}$, and $\mathbf{q}$ lie on a geodesic if the determinant $\Delta$,

$$\Delta \equiv \begin{vmatrix} (p^{01})^{1/2} & (p^{02})^{1/2} & (1-p^{01}-p^{02})^{1/2} \\ (p^1)^{1/2} & (p^2)^{1/2} & (1-p^1-p^2)^{1/2} \\ (q^1)^{1/2} & (q^2)^{1/2} & (1-q^1-q^2)^{1/2} \end{vmatrix} \tag{21}$$

vanishes, $\Delta = 0$. Now, using (3),

$$\Delta^2 = \begin{vmatrix} 1 & \cos d_{01} & \cos d_{02} \\ \cos d_{01} & 1 & \cos d_{12} \\ \cos d_{02} & \cos d_{12} & 1 \end{vmatrix} \tag{22}$$

The solution of $\Delta^2 = 0$ is $d_{02} = d_{01} \pm d_{12}$. ▌

### 3.3. General Displacements: The Cosine Law

In the general case, where $\delta\mathbf{p}$ and $\delta\mathbf{q}$ are neither orthogonal nor parallel to one another, we can interpret the angle between $\delta\mathbf{p}$ and $\delta\mathbf{q}$ in terms of the law of cosines. Indeed one readily verifies that in the general case (cf. Fig. 2)

$$\cos d_{02} = \cos d_{01} \cos d_{12} + \sin d_{01} \sin d_{02} \cos \gamma \tag{23}$$

where $\cos\gamma$ is the angle between $\delta\mathbf{p}$ and $\delta\mathbf{q}$

$$\cos\gamma = \delta\mathbf{p} \cdot \delta\mathbf{q} / [(\delta\mathbf{p} \cdot \delta\mathbf{p})(\delta\mathbf{q} \cdot \delta\mathbf{q})]^{1/2} \tag{24}$$

The result (23) will be recognized as the law of cosines for spherical triangles (cf. Fig. 2).

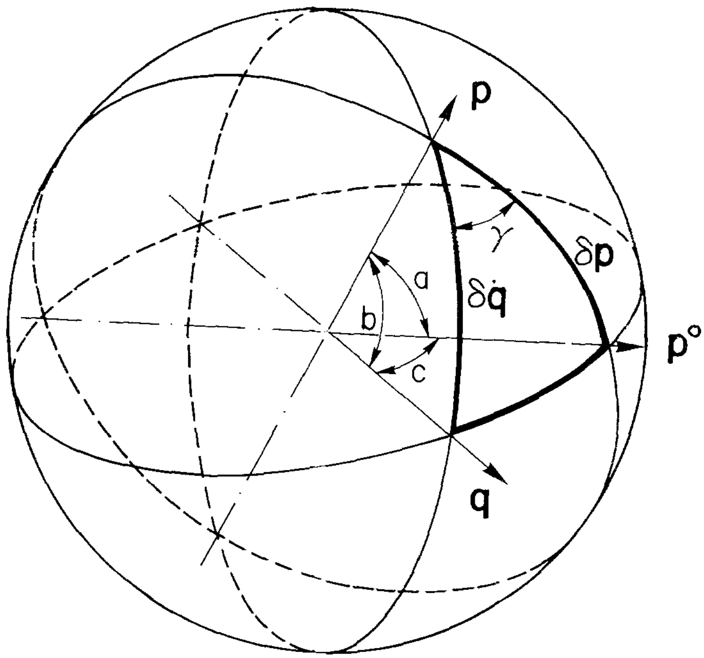


Fig. 2. The spherical triangle defined by $\mathbf{p}^0$, $\mathbf{p}$, and $\mathbf{q}$ as its three vertices. Shown are the three great circles such that any two distributions lie on one circle. $c = d_{02}$, $a = d_{01}$, $b = d_{12}$.

## 4. CONCLUDING REMARKS

Geometrical concepts can be used to advantage in interpreting the statistical description of systems. An example discussed in some detail is that the removal of a constraint corresponds to an orthogonal projection.[3] It is our intention and hope that such considerations will prove useful in discussing processes which occur at a finite rate and are thus accompanied by dissipation.

## ACKNOWLEDGMENT

We thank the U.S.–Israel Binational Science Foundation (BSF) Jerusalem, Israel, and the Petroleum Research Fund administered by the American Chemical Society for support of this work. We thank Dr. M. Berman for computational assistance.

[3] After this paper was submitted we were informed by Prof. Y. Alhassid that together with Prof. R. Balian they have obtained the same results.

## REFERENCES


1. F. Weinhold, *J. Chem. Phys.* **63**:2479, 2484, 2488, 2496 (1975); **65**:559 (1976); *Acc. Chem. Res.* **9**:236 (1976).
2. P. Salamon and R. S. Berry, *Phys. Rev. Lett.* **51**:1127 (1983).
3. J. D. Nulton, P. Salamon, B. Andresen, and Qi Anmin, *J. Chem. Phys.* **83**:334 (1985).
4. G. Ruppeiner, *Phys. Rev.* **A20**:1608 (1979); **24**:488 (1981).
5. G. Ruppeiner, *Phys. Rev. Lett.* **50**:287 (1983); *Phys. Rev.* **A27**:1116 (1983).
6. P. Salamon, J. D. Nulton, and R. S. Berry, *J. Chem. Phys.* **82**:2433 (1985); P. Salamon, J. D. Nulton, and E. Ihrig, *J. Chem. Phys.* **80**:436 (1984).
7. R. D. Levine, Geometry in classical statistical thermodynamics, *J. Chem. Phys.* (to be published).
8. W. K. Wootters, *Phys. Rev.* **D23**:357 (1981).
9. I. Procaccia and R. D. Levine, *J. Chem. Phys.* **65**:3356 (1976).
10. F. Schlögl, *Phys. Rep.* **62**:267 (1980).
11. R. D. Levine and M. Tribus, eds., *Maximum Entropy Formalism* (M.I.T. Press, Cambridge, Massachusetts, 1979).
12. A. Hobson and B.-K. Cheng, *J. Stat. Phys.* **7**:301 (1973).
13. R. D. Levine, *J. Chem. Phys.* **65**:3302 (1976).
14. I. Procaccia, Y. Shimoni, and R. D. Levine, *J. Chem. Phys.* **65**:3284 (1976).