Proc. Nat. Acad. Sci. USA
Vol. 71, No. 8, pp. 3050–3053, August 1974

Classical Statistical Mechanics of Constraints: A Theorem and Application to Polymers

(freely jointed chain/metric determinant)

MARSHALL FIXMAN

Department of Chemistry, Yale University, New Haven, Connecticut 06520

Contributed by Marshall Fixman, June 3, 1974

ABSTRACT **A classical system of mass points subject to holonomic constraints has a kinetic energy dependent on the coordinates as well as the momenta of the remaining degrees of freedom. Coordinate averages formed in the reduced space of unconstrained coordinates and their conjugate momenta then involve a metric determinant that may be difficult to evaluate. A theorem is derived that permits a relatively easy evaluation when the constraints are distances between particles, and an application is made to a Kramers type freely jointed chain.**

Background

In most formulations of the statistical mechanics of polymer chains it is supposed that to a sufficient approximation the bond lengths and angles may be constrained to constant values, and the rotational angles treated classically. One may further suppose, as did Kramers (1) and apparently Kirkwood (2–5;*), that (*1*): the momenta conjugate to the constrained coordinates are ignorable. If they are ignored, then an additional coupling arises between the unconstrained coordinates, a coupling not present in the potential energy. In recent calculations Hassager (6) has adopted this model for the freely jointed chain. *Alternatively* one could suppose that (*2*): all conjugate momenta are thermally activated. Then at equilibrium the coupling between unconstrained coordinates arises solely from the potential energy. If only bond lengths are constrained, model 2 reduces to the conventional "freely jointed chain," for which the bond vectors are independently oriented.

To avoid needless confusion, we will refer to model 1 as a Kramers chain, rather than a freely jointed chain, and we urge that others do likewise. The chain with constrained bond angles and constrained conjugate momenta, also a model 1, will be called a Kramers-Kirkwood chain.

A question as to the correctness of one or the other models 1 or 2 may naturally be approached from several points of view. (*a*) Physical (Gō and Scheraga, ref. 7): Can the "important" vibrational degrees of freedom of a chain, possibly loaded by the solvent inertia, be treated classically, and if so, more accurately by model 1 or 2? (*b*) Theoretical (Brillouin, ref. 8): Given the acceptability of a classical formulation, is it proper as in model 1 to set the velocities of constrained coordinates identically zero in the Lagrangian, not merely along a natural trajectory, about which there is no question, but also prior to the construction of conjugate momenta? As Brillouin has noted (8), momenta conjugate to constrained coordinates will ordinarily receive contributions from the velocities of unconstrained coordinates, and will not vanish†.

For the present these interesting questions will be ignored and a purely technical problem encountered with model 1 will be considered, namely the evaluation of the metric determinant. Kramers found simple formulas for short "Kramers" chains, but was not able to give a general expression. We present one here.

The problem

A system of $N + 1$ mass points is subjected to r holonomic constraints $(b) = (b_1,\ldots,b_r)$,

$$b_i = b_i(\mathbf{R}_0,\mathbf{R}_1,\mathbf{R}_2,\ldots,\mathbf{R}_N) = \text{constant};\ i = 1,2,\ldots,r \quad [1]$$

where $\mathbf{R}_i$ is the vector position of particle i in ordinary three-dimensional space. Choose the remaining f degrees of freedom to be $(a) = (a_1,a_2,\ldots,a_f)$, so that $f + r = 3(N + 1)$, and

$$\mathbf{R}_i = \mathbf{R}_i(a,b) \quad [2]$$

Given a function $F(a)$, *and* the assumption that the momenta conjugate to the b's are zero, or constant, or for some other reason ignorable, the equilibrium average of $F(a)$ is (9), for model 1,

$$\langle F(a)\rangle_\alpha = \int F(a)e^{-U(a)}[g^\alpha(a)]^{1/2}d(a) \Big/ \int e^{-U(a)} \times [g^\alpha(a)]^{1/2}d(a) \quad [3a]$$

where $g^\alpha(a)$ is the determinant of a metric tensor $G_{ij}{}^\alpha$ and

* It seems difficult, not to say impossible, to recognize the distinction between models 1 and 2 in the papers of Kirkwood and his collaborators. Our belief is that in the papers of refs. 4 and 5, and earlier, Kirkwood was using model 1 (no constraining forces appear, and g^α appears in the proper way for model 1). But it was then recognized by Erpenbeck and Kirkwood that hydrodynamic interaction was improperly treated, basically because of hydrodynamic coupling between what we have called α and β space, and revised diffusion equations were written for the full $(\alpha + \beta)$ space. At first these equations omitted constraining forces, but they were later supplied in the Erratum listed in ref. 3. So the final version contains the metric determinant g rather than g^α, and has explicit constraining forces, and, is, therefore, a class 2 model.

† A rather subtle danger should be mentioned. The kinetic energy metric tensor will ordinarily *not* be separable into blocks containing no cross terms between constrained and unconstrained variables. The artificial orthogonalization we use below is handy for our purposes, but could not be used for the construction of conjugate velocities or momenta. Their definitions require that differentials of coordinates be paramaterized in time, i.e., that they be dynamical rather than arbitrary variations.

$U(a)$ is a potential energy measured in units of kT. The tensor $G_{ij}{}^{\alpha}$ is well known, and its explicit form is considered below. The *problem* is to evaluate its determinant. We proceed without necessarily subscribing to the assumptions that make Eq. **3a** valid.

For model 2 the proper average is

$$\langle F(a)\rangle = Z^{-1}\int F(a)\exp[-U(a) - V(b)]\prod_0^N d\mathbf{R}_i \quad [3b]$$

where Z is an appropriate normalizing constant, and $V(b)$ is a potential which guarantees the satisfaction of constraints. Eq. **3b** may be brought into the form of Eq. **3a**, with g^{α} replaced by $g(a,b)$, g being the metric tensor in the full (a,b) space. As is noted in more detail below, g is essentially trivial for polymer problems. Our focus, then, is on g^{α}.

VARIOUS METRIC TENSORS

The kinetic energy in phase space

The kinetic energy T is

$$T = \frac{1}{2}\sum_{k=0}^{N} m_k|\dot{\mathbf{R}}_k|^2 = \sum_0^N |\dot{\varrho}_k|^2 \quad [4]$$

where ϱ_k is a mass weighted vector position $\varrho_k = \mathbf{R}_k\sqrt{m_k/2}$. Then (9)

$$G_{ij}{}^{\alpha} = \frac{\partial\varrho^{\alpha}}{\partial a_i}\cdot\frac{\partial\varrho^{\alpha}}{\partial a_j} \quad [5]$$

where $d\varrho^{\alpha}$ is an infinitesimal vector in the α space (a), of f dimensions. The superscript is not necessary, since

$$\frac{\partial\varrho^{\alpha}}{\partial a_i} = \frac{\partial\varrho}{\partial a_i} \quad [6]$$

but it will be helpful to keep relations like Eq. **6** explicit. Here $d\varrho$ is an infinitesimal vector in the full (a,b) space, but the partial derivative chooses $d\varrho = d\varrho^{\alpha}$.

We would like to consider the full space $(c) = (a,b)$ in more detail, because it will often be possible to evaluate the metric determinant g in the full space by simple means and this will help to evaluate g^{α}. This simplicity arises especially when the r b's are magnitudes of vectors joining two particles, and $2r$ a's are associated angles. Let $d\tau$ be an element of volume in the full ρ space.

$$d\tau = \prod_0^N d\varrho_k = \prod_0^N d\mathbf{R}_k(m_k/2)^{1/2} \quad [7]$$

$$= \sqrt{g}\, d(c) \quad [8]$$

We will suppose that g is known by some essentially trivial calculation, as in the polymer application to follow. Otherwise the theorem to be derived will not be very useful. Obviously one would not wish to study it by frontal attack, from $g = |G_{ij}|$ with G_{ij} defined by simple generalization of (**5**):

$$G_{ij} = \frac{\partial\varrho}{\partial c_i}\cdot\frac{\partial\varrho}{\partial c_j};\ i,j = 1,2,\ldots,3(N+1) \quad [9]$$

This procedure would give the correct g, but would be harder than the original problem if any c_i were simply one of the a's or b's. Instead g will be determined from the Jacobian of a vector transformation.

Considering g to be known, we proceed to an alternative evaluation that relates g to g^{α} and a more tractable metric determinant.

An orthogonal space

In general the vectors $\partial\varrho/\partial c_i$, with c_i one of the a's or b's, will not form an orthogonal set, that is the G_{ij} of (**9**) will be non-vanishing, for practically all pairs of i and j, and particularly for $c_i = a_i$, $c_j = b_j$ (6). The vectors $\partial\varrho/\partial c_i$ are the covariant basis vectors, and define a curvilinear set of axes. In order to span the full space and generate an alternative expression for g, we decompose an arbitrary displacement $d\varrho$ into

$$d\varrho = d\varrho^{\alpha} + d\varrho^{\beta}; \quad [10]$$

where

$$d\varrho^{\alpha}\cdot d\varrho^{\beta} \equiv 0 \quad [11]$$

$$d\varrho^{\alpha} = \sum_1^f (\partial\varrho^{\alpha}/\partial a_i)da_i \quad [12]$$

$$d\varrho^{\beta} = \sum_1^r (\partial\varrho^{\beta}/\partial b_i)db_i \quad [13]$$

Here $\partial\varrho^{\alpha}/\partial a_i$ is the vector previously introduced; it satisfies Eq. **6**. But

$$\partial\varrho^{\beta}/\partial b_i \neq \partial\varrho/\partial b_i \quad [14]$$

Rather $d\varrho^{\beta}$ is chosen to satisfy (**11**). The orthogonality can be achieved in principle inasmuch as the r unit vectors required to span β space can be chosen to lie along the r vectors $\partial b_i/\partial\varrho^{\beta}$. These vectors are all perpendicular to surfaces of constant b's, and, therefore, they are all perpendicular to $d\varrho^{\alpha}$, though not to each other. And most importantly,

$$\partial b_i/\partial\varrho^{\beta} = (\partial b_i/\partial\varrho^{\alpha}) + (\partial b_i/\partial\varrho^{\beta}) = (\partial b_i/\partial\varrho) \quad [15]$$

since the definition of $d\varrho^{\alpha}$ makes $\partial b_i/\partial\varrho^{\alpha} = 0$. (A minor technical point might be noted; Eq. **15** implies that all vectors present have the same dimensionality; they can be spaced out with zero elements to give a dimensionality of $3(N+1)$ as necessary.) Note also that $\partial\varrho^{\beta}/\partial b_i$ is not a derivative at constant (a); as b_i changes, the a's have to change in such a way as to maintain the orthogonality.

It follows from the orthogonality that

$$|d\varrho|^2 = \sum_1^f\sum_1^f G_{ij}{}^{\alpha}da_ida_j + \sum_1^r\sum_1^r G_{ij}{}^{\beta}db_idb_j \quad [16]$$

with $G_{ij}{}^{\alpha}$ as defined before, and

$$G_{ij}{}^{\beta} = (\partial\varrho^{\beta}/\partial b_i)\cdot(\partial\varrho^{\beta}/\partial b_j);\ i,j = 1,2,\ldots,r$$

$$= \sum_{k=1}^{r}(\partial\varrho_k{}^{\beta}/\partial b_i)(\partial\varrho_k{}^{\beta}/\partial b_j) \quad [17]$$

if we choose r mutually orthogonal unit vectors to span the space generated by $\partial b_i/\partial\varrho^{\beta}$, and take $d\varrho_k{}^{\beta}$ to be the projection of $d\varrho^{\beta}$ along the kth such unit vector.

Eq. **16** implies that (verification is given in the *Appendix*)

$$g = g^{\alpha}g^{\beta}$$

where g^{β} is the determinant of $G_{ij}{}^{\beta}$, or

$$g^{\alpha} = g/g^{\beta} = gh \quad [18]$$

where h is the determinant of H_{ij}, this being the matrix inverse of G_{ij}^{β},

$$H = (G^{\beta})^{-1}$$

$$H_{ij} = \sum_{k=1}^{r} (\partial b_i/\partial \varrho^{\beta}{}_k)(\partial b_j/\partial \varrho_k{}^{\beta})$$

Eq. **15** gives

$$H_{ij} = (\partial b_i/\partial \varrho^{\beta}) \cdot (\partial b_j/\partial \varrho^{\beta}) = (\partial b_i/\partial \varrho) \cdot (\partial b_j/\partial \varrho)$$

$$= \sum_{l=0}^{N} (2/m_l)(\partial b_i/\partial \mathbf{R}_l) \cdot (\partial b_j/\partial \mathbf{R}_l) \quad [19]$$

Eqs. **18** and **19** constitute the theorem in its general form. The familiar reduction of a microcanonical average to an integral over a constant energy surface is a special case of Eq. **18**.

APPLICATION TO CHAIN POLYMERS

A polymer chain of $N + 1$ beads numbered from 0 to N is taken to have fixed values of

$$b_i = |\mathbf{b}_i| = |\mathbf{R}_i - \mathbf{R}_{i-1}|, \; i = 1,2,\ldots,N, \quad [20]$$

and the $2N + 3$ remaining coordinates (a) are taken to be the center of mass $\mathbf{R}_c$ and the polar angles θ_i, ϕ_i of bond vector $\mathbf{b}_i$ in a laboratory frame (or in a frame defined relative to preceding bond vectors; the alternative forms for g^{α} differ trivially). Now

$$\prod_0^N d\mathbf{R}_i \propto d\mathbf{R}_c \prod_1^N d\mathbf{b}_i \propto d\mathbf{R}_c S \prod_1^N d\theta_i d\varphi_i db_i \quad [21]$$

where

$$S \equiv \prod_1^N \sin\theta_i \quad [22]$$

and we omit any proportionality factors (numerical constants, Πb_i^2) not dependent on chain configuration. It follows from Eqs. **7** and **8** that

$$\sqrt{g} = S \quad [23]$$

If $\mathbf{b}_i$ for $i \geqslant 2$ is expressed in internal coordinates, θ_i for $i \geqslant 2$ becomes equal to the supplement to the bond angle between $\mathbf{b}_i$ and $\mathbf{b}_{i-1}$, but θ_1 refers still to a lab frame.

Now Eq. **20** gives

$$\partial b_i/\partial \mathbf{R}_k = \mathbf{e}_i(\delta_{ik} - \delta_{i-1,k}) \quad [24]$$

where $\mathbf{e}_i = \mathbf{b}_i/b_i$, a unit vector along $\mathbf{b}_i$. Therefore, for $i,j = 1,\ldots,N$, Eq. **19** gives

$$H_{ij} = (\mathbf{e}_i \cdot \mathbf{e}_j) \sum_{k=0}^{N} m_k{}^{-1}(\delta_{ik} - \delta_{i-1,k})(\delta_{jk} - \delta_{j-1,k}) \quad [25]$$

$$= \mathbf{e}_i \cdot \mathbf{e}_j(m_i{}^{-1}\delta_{ij} - m_j{}^{-1}\delta_{i-1,j} - m_i{}^{-1}\delta_{j-1,i} + m_{i-1}{}^{-1}\delta_{ij}) \quad [26]$$

A section of H, the tridiagonal matrix of H_{ij}, is shown below.

$$H = \begin{bmatrix} \mu_1 & \gamma_1 & 0 & \cdots & 0 \\ \gamma_1 & \mu_2 & & & \cdot \\ 0 & \cdot & & & \cdot \\ \cdot & & \cdot & & \cdot \\ \cdot & & & \cdot & 0 \\ \cdot & & & & \mu_{N-1}\gamma_{N-1} \\ 0 & \cdots & 0 & \gamma_{N-1} & \mu_N \end{bmatrix} \quad [27]$$

where

$$\mu_i = (1/m_{i-1}) + (1/m_i) \quad [28]$$

$$\gamma_i = (\cos\chi_i)/m_i = -\mathbf{e}_i \cdot \mathbf{e}_{i+1}/m_i \quad [29]$$

and χ_i is the bond angle. Eqs. **18** and **23** give

$$g^{\alpha} = \left[\prod_1^N \sin\theta_i\right]^2 h \quad [30]$$

For the evaluation of h_N, the determinant of H shown in (**27**) for N bonds, a recursion relation is easily derived by expansion in minors:

$$h_N = \mu_N h_{N-1} - (\gamma_{N-1})^2 h_{N-2}; \; h_1 = \mu_1; \; h_0 = 1. \quad [31]$$

Formulas for small N are easily derived by iteration. The results agree for small N with the h_N given by Kramers (1) and Hassager (6), who apparently used more complex methods. It is interesting that h_N does not depend on the bond lengths.

It seems probable that the important regions of phase space will lie near $\gamma_i = 0$ for equilibrium, or near some other constant values in other problems. Then an elementary perturbation theory of (**31**) may be useful. For example, consider the chain of uniform mass elements. Without loss in generality we take all $\mu_i = 1$ (all $m_i = 2$). Then let

$$h_N/h_{N-1} = \exp(\delta_N); \; \delta_1 = 0 \quad [32]$$

$$h_N = \exp\left(\sum_2^N \delta_k\right) \quad [33]$$

Eq. **31** gives‡

$$\exp(\delta_N) = 1 - (\gamma_{N-1})^2 \exp(-\delta_{N-1}) \quad [34]$$

$$= 1 - (\gamma_{N-1})^2[1 + \gamma_{N-2}{}^2 + \ldots]$$

or

$$\delta_N = -\gamma_{N-1}{}^2\left[1 + \left(\gamma_{N-2}{}^2 + \frac{1}{2}\gamma_{N-1}{}^2\right) + \ldots\right]$$

or

$$h_N = \exp\left[-\frac{1}{4}\sum_1^{N-1} \cos^2 x_i + \ldots\right] \quad [35]$$

Although h_N may stray far from unity, its effect on the angular distribution of any particular bond vector appears likely to be small.

DISCUSSION

Imposition of constraints on the bond angles would lead to a somewhat more difficult problem. Probably h would be easiest to evaluate if the degrees of freedom were chosen to be the bond rotational angles ϕ_i, and $2N - 1$ constrained coordinates $|\mathbf{b}_i|$ and $|\mathbf{b}_i + \mathbf{b}_{i+1}|$ were chosen to fill out coordinate phase space.

The constraints, whether on bond lengths and angles or just on lengths, are clearly annoying, and our own prejudices lie heavily in favor of model 2, rather than model 1 (the Kramers–Kirkwood models of the introduction). It has to be said, nevertheless, that metric tensors of the form H arise also in the nonequilibrium theory (M. Fixman and J. Kovac, manu-

‡ The recursion relations with $\gamma_k{}^2 \leqslant 1/4$ establish that $0 \geqslant \delta_k \geqslant (-\ln 2)$, as may be verified by induction. So the metric determinant is well behaved.

script in preparation) of model 2, for the constraining forces present in model 2 make vanish the mean drift velocities such as $\dot{\mathbf{b}}_i$ (in the sense of Brownian motion theory), even if the instantaneous velocities have a Maxwellian distribution. So one does not escape the necessity to deal with quantities like H. However, we have persuaded ourselves (M. Fixman and J. Kovac, manuscript in preparation) that for probabilities calculated near equilibrium, and not involving gross elongations of the chain, the H_{ij} may be reasonably approximated by their equilibrium averages. When strong external forces are present, the appropriate nonequilibrium average of H_{ij} could be used. In either situation, whenever fluctuations of H, and therefore also of h, are ignored, the constant value of h may be cancelled from equations for averages, and the distinction between models 1 and 2 is erased. The quantitative results derived herein for a freely jointed chain, for which fluctuations are likely to be larger than for stiff chains, encourages us to regard the distinction as having little practical consequence. Hassager (6) has already made this point on the basis of detailed numerical calculations for short chains.

This work was supported in part by U.S. National Institutes of Health Grant GM 13556.

APPENDIX (PROOF OF EQ. 18)

Let the complete covariant metric tensor be partitioned into blocks:

$$G = \begin{pmatrix} G^{\alpha} & G^{\alpha b} \\ G^{b\alpha} & G^{b} \end{pmatrix}; \; G_{ij} = (\partial \varrho/\partial c_i) \cdot (\partial \varrho/\partial c_j), \qquad [\mathbf{A1}]$$

where c_i is selected from the set (a) in the first row, and from (b) in the second. Likewise c_j is selected from (a) in the first column, and from (b) in the second. Thus G^{α} is the square matrix of dimension f, the determinant of which is wanted. Similarly partition the contravariant tensor

$$G^{-1} = \begin{pmatrix} H^{a} & H^{a\beta} \\ H^{\beta a} & H \end{pmatrix}; \; H_{ij} = (\partial c_i/\partial \varrho) \cdot (\partial c_j/\partial \varrho) \qquad [\mathbf{A2}]$$

where H is the square matrix of dimension r, the determinant of which is (relatively) easy to compute. The requirement that

$$GG^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad [\mathbf{A3}]$$

imposes conditions on the submatrices of G and G^{-1} that permit an immediate verification of

$$G \begin{pmatrix} 1 & H^{a\beta} \\ 0 & H \end{pmatrix} = \begin{pmatrix} G^{\alpha} & 0 \\ G^{b\alpha} & 1 \end{pmatrix}. \qquad [\mathbf{A4}]$$

Calculation of the determinant of both sides gives Eq. **18**.

1. Kramers, H. A. (1946) *J. Chem. Phys.* **14**, 415–424.
2. Erpenbeck, J. J. & Kirkwood, J. G. (1958) *J. Chem. Phys.* **29**, 909–913.
3. Erpenbeck, J. J. & Kirkwood, J. G. (1963) *J. Chem. Phys.* **38**, 1023–1024.
4. Kirkwood, J. G. (1949) *Rec. Trav. Chim. Pays-Bas* **68**, 649–660.
5. Kirkwood, J. G. (1954) *J. Polym. Sci.* **12**, 1–14.
6. Hassager, O. (1974) *J. Chem. Phys.* **60**, 2111–2124.
7. Gō, N. & Scheraga, H. A. (1969) *J. Chem. Phys.* **51**, 4751–4767.
8. Brillouin, L. (1964) *Tensors in Mechanics and Elasticity* (Academic Press, New York), p. 231.
9. Gibbs, J. W. (1948) *Collected Works* (Yale University Press, New Haven, Conn.), Vol. II.