JOURNAL OF COMPUTATIONAL PHYSICS 58, 409–416 (1985)

Statistical Mechanics of Hard Ellipsoids. I. Overlap Algorithm and the Contact Function

JOHN W. PERRAM*

Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, New York 10012

AND

M. S. WERTHEIM†

Mathematics Department, Rutgers University, New Brunswick, New Jersey 08903

Received February 9, 1984

A contact function for two arbitrary ellipsoids is derived. The numerical value of the contact function is less than 1 if they overlap, and greater than 1 if they do not. This extends previous work by Vieillard-Baron, who derived an overlap criterion for spheroids without use of a contact function. The equivalence of the two criteria has been checked by extensive numerical test with spheroids. It is shown that the use of the contact function greatly facilitates the calculation of the pressure in Monte Carlo simulations. ©1985 Academic Press, Inc.

1. INTRODUCTION

The modeling of molecules by hard bodies plays a key role in statistical theories, because the spatial exclusion due to harsh repulsive forces is the dominant factor in determining the fluid structure as measured by the pair correlation function. For atomic fluids, the idea of a reference system of hard spheres appeared long ago in the work of van der Waals [1], and now forms the basis of quantitatively successful perturbation theories [2, 3].

In choosing hard core shapes for modeling molecular fluids the following criteria are relevant:

(a) degree of approximation to the actual repulsion,

(b) generality,

(c) required number of geometric parameters,

(d) ease of deciding whether or not two cores overlap.

* Permanent address: Matematisk Institut, Odense University, Campusvej 55, 5230 Odense M, Denmark.

† Address for correspondence.

0021-9991/85 $3.00

Most existing calculations are for cores composed of two fused hard spheres [4–6], which are the simplest model with respect to criterion (d). Other calculations use spherocylinders [7, 8], where the overlap decision is somewhat harder, but still tractable. The model of hard ellipsoids is excellent with respect to criteria (a), (b) and (c). The difficulty of deciding overlap has so far discouraged its use.

In this paper we derive a numerically convenient algorithm for deciding overlap of two ellipsoids A and B. This is done by constructing a positive function $F_{AB}(1, 2)$ with the following property:

$$F_{AB}(1,2) \left\{ \begin{matrix} <1 \\ =1 \\ >1 \end{matrix} \right\} \text{ if } A \text{ and } B \left\{ \begin{matrix} \text{overlap} \\ \text{are externally tangent} \\ \text{do not overlap.} \end{matrix} \right\}. \tag{1.1}$$

This improves on the overlap crriterion for spheroids derived by Vieillard-Baron [9]. The latter result is not of the form (1.1) and therefore is inconvenient in simulation studies.

2. One Ellipsoid

The desired contact function $F_{AB}(1, 2)$ is obtained with the aid of a function $F_A(\mathbf{r}-\mathbf{r}_A, \Omega_A)$, which refers to a single ellipsoid A, specified by the location $\mathbf{r}_A$ of its center, and the angles Ω_A which express the orientation in space. The function F_A is required to be non-negative and to satisfy

$$F_A(\mathbf{r}-\mathbf{r}_A, \Omega_A) \left\{ \begin{matrix} <1 \\ =1 \\ >1 \end{matrix} \right\} \text{ for } \mathbf{r} \left\{ \begin{matrix} \text{inside } A \\ \text{on the surface of } A \\ \text{outside } A \end{matrix} \right\}. \tag{2.1}$$

The choice of F_A is not unique, but an obvious simplest choice suggests itself, namely,

$$F_A(\mathbf{r}-\mathbf{r}_A, \Omega_A) = (\mathbf{r}-\mathbf{r}_A)^{\mathrm{T}} \mathbf{A}^{-1} (\mathbf{r}-\mathbf{r}_A), \tag{2.2}$$

where T indicates the transpose, $\mathbf{r}_A$ is the center of the ellipsoid, and $\mathbf{A}$ is the matrix

$$\mathbf{A}(\Omega_A) = \sum_{i=1}^{3} \mathbf{R}_i(\Omega_A)\, \mathbf{R}_i^{\mathrm{T}}(\Omega_A), \tag{2.3}$$

where the $\mathbf{R}_i$ are the semiaxis vectors. Corresponding to the $\mathbf{R}_i$ we define unit vectors $\mathbf{u}_i = \mathbf{R}_i / R_i$.

In practice these unit vectors are specified by applying a cartesian rotation matrix $S(\Omega_A)$ to the space fixed-unit vectors,

$$\mathbf{u}_i(\Omega_A) = \mathbf{S}(\Omega_A)\, \mathbf{e}_i, \tag{2.4}$$

and therefore

$$\mathbf{A}(\Omega_A) = \mathbf{S}(\Omega_A)\,\hat{\mathbf{A}}S^{\mathrm{T}}(\Omega_A), \qquad \hat{\mathbf{A}} = \sum_{i=1}^{3} R_i^2\,\mathbf{e}_i\mathbf{e}_i^{\mathrm{T}}. \tag{2.5}$$

$\mathbf{A}$ and $\mathbf{A}^{-1}$ are symmetric and positive definite.

3. Overlap of Two Ellipsoids

We define a function

$$F(\mathbf{r}, \lambda) = \lambda F_A(\mathbf{r}) + (1-\lambda)\,F_B(\mathbf{r}), \tag{3.1}$$

which depends on the positions $\mathbf{r}_A$, $\mathbf{r}_B$ and orientations Ω_A, Ω_B of *two* ellipsoids A and B. The dependence on these parameters has been suppressed in (3.1), since we hold them fixed in studying the dependence of $F(\mathbf{r}, \lambda)$ on $\mathbf{r}$ and λ.

In all of the following the parameter λ is restricted to the interval $0 \leqslant \lambda \leqslant 1$, so that we have $F(\mathbf{r}, \lambda) \geqslant 0$. For fixed λ, the function $F(\mathbf{r}, \lambda)$ has a unique minimum as a function of $\mathbf{r}$. For $\lambda = 0$ the minimum $F = 0$ occurs at $\mathbf{r} = \mathbf{r}_B$, and for $\lambda = 1$ it is $F = 0$ at $\mathbf{r} = \mathbf{r}_A$. For intermediate values of λ, the location of the minimum of F is determined by

$$\nabla F(\mathbf{r}, \lambda) = 0, \tag{3.2}$$

or explicitly

$$\lambda\mathbf{A}^{-1}(\mathbf{r}-\mathbf{r}_A) + (1-\lambda)\,\mathbf{B}^{-1}(\mathbf{r}-\mathbf{r}_B) = 0. \tag{3.3}$$

The solution $\mathbf{r}(\lambda)$ is expressed conveniently in terms of the matrix $\mathbf{C}$ defined by

$$\mathbf{C}(\lambda) = [\lambda\mathbf{B} + (1-\lambda)\,\mathbf{A}]^{-1}. \tag{3.4}$$

The existence of the inverse taken in (3.4) is guaranteed by the positive definitive property of $\mathbf{A}$ and $\mathbf{B}$. Alternative forms of the solutions of (3.3) are

$$\mathbf{r}(\lambda) - \mathbf{r}_A = (1-\lambda)\,\mathbf{A}\mathbf{C}\mathbf{r}_{AB}, \tag{3.5}$$

$$\mathbf{r}(\lambda) - \mathbf{r}_B = -\lambda\mathbf{B}\mathbf{C}\mathbf{r}_{AB}, \tag{3.6}$$

where $\mathbf{r}_{AB} = \mathbf{r}_B - \mathbf{r}_A$. Insertion of this solution in (3.1) and use of (3.4) yields an equation for $F[\mathbf{r}(\lambda), \lambda]$,

$$F[\mathbf{r}(\lambda), \lambda] = \lambda(1-\lambda)\,\mathbf{r}_{AB}^{\mathrm{T}}\,\mathbf{C}\mathbf{r}_{AB}. \tag{3.7}$$

This convenient form does not contain $\mathbf{r}(\lambda)$ explicitly.

Our quest for a contact function requires knowledge of the behavior of $F[\mathbf{r}(\lambda), \lambda]$ as λ varies from 0 to 1. This in turn calls for knowledge of the path traced out by $\mathbf{r}(\lambda)$ in the process.

In the region outside both ellipsoids we have $F(\mathbf{r}, \lambda) > 1$. If A and B do not overlap, then the path of the minimum must certainly traverse this region, so that $F[\mathbf{r}(\lambda), \lambda]$ attains values greater than 1 in $0 < \lambda < 1$. If A and B overlap, then we have $F(\mathbf{r}, \lambda) < 1$ in the shared region $A \cap B$ for any value of λ in $0 \leqslant \lambda \leqslant 1$. Hence, the minimum value $F[\mathbf{r}(\lambda), \lambda]$ is certainly less than 1. This means that the path $\mathbf{r}(\lambda)$ cannot enter the region outside both A and B, and therefore must traverse $A \cap B$.

These considerations suffice to show that we have

$$\max_{0<\lambda<1} F[\mathbf{r}(\lambda), \lambda] \begin{Bmatrix} <1 \\ =1 \\ >1 \end{Bmatrix} \text{ for } A \text{ and } B \begin{Bmatrix} \text{overlapping} \\ \text{externally tangent} \\ \text{non-overlapping} \end{Bmatrix}. \tag{3.8}$$

In order to declare this maximum a suitable contact function we need only to show that the maximum in λ is unique.

The condition for a maximum as a function of λ is

$$\frac{dF[\mathbf{r}(\lambda), \lambda]}{d\lambda} = \left[\frac{\partial F(\mathbf{r}, \lambda)}{\partial \lambda} + \mathbf{r}'(\lambda) \cdot \nabla F(\mathbf{r}, \lambda) \right]_{\mathbf{r} = \mathbf{r}(\lambda)} = 0, \tag{3.9}$$

where $\mathbf{r}'(\lambda) = d\mathbf{r}(\lambda)/d\lambda$. In view of (3.2) this becomes

$$F_A[\mathbf{r}(\lambda), \lambda] - F_B[\mathbf{r}(\lambda), \lambda] = 0. \tag{3.10}$$

The geometric interpretation of (3.10) is noted later. In order to establish the uniqueness of the maximum, we compute the second derivative,

$$\frac{d^2F[\mathbf{r}(\lambda), \lambda]}{d\lambda^2} = 2[(\mathbf{r} - \mathbf{r}_A)^{\mathrm{T}} \mathbf{A}^{-1} - (\mathbf{r} - \mathbf{r}_B)^{\mathrm{T}} \mathbf{B}^{-1}] \mathbf{r}'(\lambda). \tag{3.11}$$

By the use of (3.4) and (3.5) this becomes

$$\frac{d^2F[\mathbf{r}(\lambda), \lambda]}{d\lambda^2} = 2\mathbf{r}_{AB}^{\mathrm{T}} \mathbf{C}(\lambda) \mathbf{r}'(\lambda). \tag{3.12}$$

Differentiation of (3.3) yields a relation for $\mathbf{r}'(\lambda)$,

$$[\lambda \mathbf{A}^{-1} + (1 - \lambda) \mathbf{B}^{-1}] \mathbf{r}'(\lambda) = -\mathbf{C}(\lambda) \mathbf{r}_{AB}. \tag{3.13}$$

Elimination of $\mathbf{r}'(\lambda)$ between (3.12) and (3.13) leads to the explicit result

$$\frac{d^2F[\mathbf{r}(\lambda), \lambda]}{d\lambda^2} = -2\mathbf{r}_{AB}^{\mathrm{T}} \mathbf{C}(\lambda) [\lambda \mathbf{A}^{-1} + (1 - \lambda) \mathbf{B}^{-1}] \mathbf{C}(\lambda) \mathbf{r}_{AB}, \tag{3.14}$$

which is negative definite. The concavity of $F[\mathbf{r}(\lambda), \lambda]$ in λ implies the uniqueness of the maximum in the interval $0 \leqslant \lambda \leqslant 1$. Therefore, the desired contact function F_{AB} is

$$F_{AB}(\mathbf{r}_{AB}, \Omega_A, \Omega_B) = \max_{0 \leqslant \lambda \leqslant 1} F[\mathbf{r}(\lambda), \lambda]. \tag{3.15}$$

The contact function has a geometric interpretation in terms of ellipsoid scaling, defined as expansion or contraction of all linear distances of the ellipsoid by a positive scale factor μ, while holding the position of the center and the orientation in space constant. For $i = A$ or B, $F_i(\mathbf{r}) = \mu_i^2$ represents a scaled ellipsoid with linear scale factor μ_i. Since $\nabla F_i(\mathbf{r})$ is in the direction of the normal to the scaled ellipsoid through $\mathbf{r}$, the path $\mathbf{r}(\lambda)$ as defined by (3.1) and (3.2) represents the locus of points of external tangency between scaled ellipsoids with linear scale factors given by

$$F_A[\mathbf{r}(\lambda), \lambda] = \mu_A^2, \qquad F_B[\mathbf{r}(\lambda), \lambda] = \mu_B^2. \tag{3.16}$$

The condition (3.10) which determines the maximum of $F[\mathbf{r}(\lambda), \lambda]$ selects the unique point on the path of $\mathbf{r}(\lambda)$ for which the scale factors are equal, $\mu_A = \mu_B$.

The end result is

$$F_{AB}(\mathbf{r}_{AB}, \Omega_A, \Omega_B) = \mu^2, \tag{3.17}$$

where μ is the common linear scale factor that must be applied to A and B in order to secure external tangency.

4. The Pressure in Terms of the Contact Function

The contact function, introduced as a criterion for overlap, plays an additional important role in the calculation of the pressure from Monto Carlo configurations.

The standard result used to obtain p is the virial equation of state, deduced by a scaling of all lengths. The periodicity of the Monte Carlo system does not affect the derivation. For the special case of hard particles we have

$$\frac{p}{kT} = \rho + \frac{1}{6} \int_{S_c} \rho^{(2)}(\mathbf{r}_{AB}, \Omega_A, \Omega_B)\, \mathbf{n} \cdot \mathbf{r}_{AB} d\mathbf{r}_{ab}\, d\Omega_A\, d\Omega_B, \tag{4.1}$$

where T is the Kelvin temperature, k Boltzmann's constant, ρ the number density, and $\rho^{(2)}$ is the non-normalized orientation-dependent pair distribution function. The surface of contact S_c is traced out by $\mathbf{r}_{AB}$ as B moves over all contact configurations with A, at fixed $\mathbf{r}_A$, Ω_A and Ω_B. The unit normal $\mathbf{n}$ is outward for A and inward for B.

In dealing with Monte Carlo configurations, the integral over the contact surface S_c must be approximated by counting near-contact configurations within a thin shell surrounding S_c. The thickness of the shell should be small enough to make the effect of $\mathbf{n} \cdot \nabla \rho^{(2)}(1, 2)$ negligible.

Away from the low-density limit, the function $\rho^{(2)}(\mathbf{r}_{AB}, \Omega_A, \Omega_B)$ at constant Ω_A and Ω_B still depends on the position vector $\mathbf{r}_{AB}$ on the contact surface $S_c(\Omega_A, \Omega_B)$. This variation of $\rho^{(2)}(1, 2)$ over S_c imposes the constraint that the proportionality of the integral over the shell to the integral over S_c must be satisfied for each direction of $\mathbf{r}_{AB}$ separately. For a shell of infinitesimal thickness $dt(\mathbf{r}_{AB})$ with weight function $w(\mathbf{r}_{AB})$, we must have

$$w(\mathbf{r}_{AB})\, dt(\mathbf{r}_{AB}) = \text{constant} \times \mathbf{r}_{AB} \cdot \mathbf{n}(\mathbf{r}_{AB}). \tag{4.2}$$

The easiest way to satisfy this is the adoption of constant w with $dt(\mathbf{r}_{AB})$ proportional to $\mathbf{r}_{AB} \cdot \mathbf{n}$. Such a shell thickness is provided in a natural way by the contact function. As a consequence of (3.16), the shell thickness between two neighboring shells

$$F_{AB}(\mathbf{r}_{AB}) = \mu^2 \qquad \text{and} \qquad F_{AB}(\mathbf{r}_{AB}) = (\mu + d\mu)^2$$

is given by

$$dt(\mathbf{r}_{AB}) = \mathbf{r}_{AB} \cdot \mathbf{n}(\mathbf{r}_{AB})\, d\mu. \tag{4.3}$$

For a shell of *finite* thickness we still obtain a very simple result, as long as neglect of the variation of $\rho^{(2)}$ in the $\mathbf{r}_{AB}$-direction is justified within the shell. Take the shell between the two surfaces $F_{AB}(\mathbf{r}_{AB}) = 1$ and $F_{AB}(\mathbf{r}_{AB}) = \mu^2$ and a narrow cone of rays surrounding a given direction $\mathbf{r}$. Let δV be the volume element that is in both the shell and the cone, and let δS_1 be the surface intercepted by the cone on $F_{AB}(\mathbf{r}_{AB}) = 1$. For an arbitrary weight function $w(\mu)$, which is independent of the position $\mathbf{r}$ on the surface of constant μ, we have

$$\int_{\delta V} w(\mu)\, d\mathbf{r} = \int_0^{\mu} w(x)\, x^2\, dx \int_{S_1} \mathbf{r} \cdot \mathbf{n}(\mathbf{r})\, dS_1. \tag{4.4}$$

Specialization of this exact result to $w(\mu) = 1$ results in the approximation

$$\int_{S_c} \rho^{(2)}(1, 2)\, \mathbf{r}_{AB} \cdot \mathbf{n} dS = \frac{3}{\mu^3 - 1} \int_{1 < F_{AB} < \mu^2} \rho^{(2)}(\mathbf{r}_{AB})\, d\mathbf{r}_{AB}. \tag{4.5}$$

For the limiting case of constant $\rho^{(2)}(1, 2)$, Eq. (4.5) is exact for any shell thickness. The error due to non-zero $\nabla\rho^{(2)}(1, 2)$ is of order $(\mu - 1)$, since both the shell thickness and the error in $\rho^{(2)}(1, 2)$ are of order $\mu - 1$.

5. The Overlap Algorithm in Simulation

The best implementation of an algorithm depends on its intended use. Our primary interest is in application to Monte Carlo simulation. Here, each step entails selecting one of the N particles, translating and rotating it by a small amount, and

then testing it for overlap with any of the other $N-1$ particles. If an overlap is detected, the move is rejected and a new particle is chosen. The size of the moves is adjusted so that about half the moves are accepted. This implies that the outcomes of the test have non-overlaps outnumbering overlaps in a ratio of roughly $3N/2$ to 1. Therefore, the overriding concern is fastest possible establishment of non-overlaps. This is accomplished by stepping toward the maximum in λ of $F[\mathbf{r}(\lambda), \lambda]$ by an iteration technique, "Localmin [10]," which finds the maximum with order of convergence 1.3247. As soon as a value $F[\mathbf{r}(\lambda), \lambda] > 1$ is found, non-overlap is established, and the search for the maximum is broken off.

A comparison of our algorithm against the overlap criterion of Vieillard-Baron [9] on 10^5 random configurations resulted in identical decisions in all cases. For over 90 percent of non-overlaps the initial guess of $\lambda = \frac{1}{2}$ resulted in decision by $F[\mathbf{r}(\frac{1}{2}), \frac{1}{2}] > 1$.

6. Computation of the Contact Function

The iterative calculation of the contact function entails calculation of $\mathbf{C}(\lambda)$ for successive values of λ. This may be done by taking the inverse in (3.4) numerically, or by using the characteristic equation to convert it into a polynomial. Any matrix $\mathbf{M}$ satisfies the equation

$$\sum_{s=0}^{d} (-1)^s T_s(\mathbf{M})\, \mathbf{M}^{d-s} = 0, \tag{6.1}$$

where d is the dimensionality, here $d=3$, $T_0 = 1$, and T_S for $s>0$ is the sum of the determinants of the $s \times s$ principal minors, of which there are $d!/s!(d-s)!$. From (6.1) we obtain

$$\mathbf{M}^{-1} = \left(\sum_{s=0}^{d-1} (-1)^{s-d-1} T_s(\mathbf{M})\, \mathbf{M}^{d-s} \right) \Big/ T_d(\mathbf{M}). \tag{6.2}$$

We are dealing with a matrix of the form $\mathbf{M} = \mathbf{G} + \mathbf{H}$, and need to rewrite the $T_s(\mathbf{M})$ in terms of the matrices $\mathbf{G}$ and $\mathbf{H}$. The results required for $d=3$ are obtained by straightforward manipulation. We find

$$T_1(\mathbf{G}+\mathbf{H}) = T_1(\mathbf{G}) + T_1(\mathbf{H}), \tag{6.3}$$

$$T_2(\mathbf{G}+\mathbf{H}) = T_2(\mathbf{G}) + T_2(\mathbf{H}) + T_1(\mathbf{G})\, T_1(\mathbf{H}) - T_1(\mathbf{GH}), \tag{6.4}$$

$$\begin{aligned} T_3(\mathbf{G}+\mathbf{H}) = {} & T_3(\mathbf{G}) + T_3(\mathbf{H}) + T_1(\mathbf{G})\, T_2(\mathbf{H}) + T_1(\mathbf{H})\, T_2(\mathbf{G}) \\ & - T_1(\mathbf{GH})[T_1(\mathbf{G}) + T_1(\mathbf{H})] + T_1(\mathbf{G}^2\mathbf{H}) + T_1(\mathbf{GH}^2). \end{aligned} \tag{6.5}$$

All of the T's can be expressed in terms of scalar products of the unit vectors $\mathbf{u}_i(A)$ and $\mathbf{u}_i(B)$.

We quote the explicit result for two identical spheroids with $R_1 = R_2$, $R_3 = eR_1$. We have

$$R_1^2 \mathbf{C}(\lambda) = \mathbf{I} + \alpha \mathbf{u}_3(A)\,\mathbf{u}_3(A)^{\mathrm{T}} + \beta \mathbf{u}_3(B)\,\mathbf{u}_3(B)^{\mathrm{T}} + \gamma[\mathbf{u}_3(A)\,\mathbf{u}_3(B)^{\mathrm{T}} + \mathbf{u}_3(B)\,\mathbf{u}_3(A)^{\mathrm{T}}], \tag{6.6}$$

where $\mathbf{I}$ is the unit matrix and

$$\alpha = x(1-y)/D, \qquad \beta = y(1-x)/D, \qquad \gamma = xy\mathbf{u}_3(A)\cdot\mathbf{u}_3(B)/D, \tag{6.7}$$

$$D = (1-x)(1-y) - xy[\mathbf{u}_3(A)\cdot\mathbf{u}_3(B)]^2, \tag{6.8}$$

$$x = (1-\lambda)(1-e^2), \qquad y = \lambda(1-e^2). \tag{6.9}$$

By using this form in (3.7), $F[r(\lambda), \lambda]$ is computed rapidly.

Acknowledgments

We acknowledge helpful discussions with S. W. DeLeeuw, J. L. Lebowitz, and E. R. Smith. J.W.P. gratefully acknowledges the hospitality of the Courant Institute of Mathematical Sciences and the use of their computers. M.S.W. acknowledges support by the U.S. Air Force under Grant AFOSR 82-0016A and by the National Science Foundation under Grant CHE-82-11236.

References

1. J. D. van der Waals, "Die Kontinuität des gasförmigen und flüssigen Zustandes," Leipzig, 1873.
2. J. A. Barker and D. Henderson, *J. Chem. Phys.* **47** (1967), 4714.
3. J. D. Weeks, D. C. Chandler, and H. C. Andersen, *J. Chem. Phys.* **54** (1971), 5237; **55** (1971), 5422. H. C. Andersen and D. Chandler, *J. Chem. Phys.* **57** (1972), 1918.
4. D. Chen and W. A. Steele, *J. Chem. Phys.* **54** (1971), 703.
5. W. Streett and D. Tildesley, *Proc. R. Soc. London Ser. A* **335** (1977), 239.
6. F. Kohler, J. W. Perram, and N. Quirke, *J. Chem. Phys.* **71** (1979), 4128. The preceding three references are a small sampling. It would be impractical to list all papers pertaining to hard atomics.
7. J. Vieillard-Baron, *Mol. Phys.* **28** (1974), 809.
8. T. Boublik, I. Nezbeda, and O. Trnka, *Czech. J. Phys. B* **26** (1976), 1081.
9. J. Vieillard-Baron, *J. Chem. Phys.* **56** (1972), 4729.
10. R. P. Brent, "Algorithms for Minimization without Derivatives," pp. 61–80, Prentice–Hall, Englewood Cliffs, N.J., 1973.