Shape Distributions for Gaussian Molecules: Circular and Linear Chains as Asymmetric Ellipsoids

Gaoyuan Wei and B. E. Eichinger*

Department of Chemistry, BG-10, University of Washington, Seattle, Washington 98195

Received January 25, 1990; Revised Manuscript Received April 26, 1990

ABSTRACT: Shape distribution functions for linear and circular macromolecules in three dimensions have been numerically evaluated on IBM 3090 supercomputers. The traditional Gaussian approximation is used to model these macromolecules, and they are found to be highly anisometric when viewed from a coordinate frame that is aligned with the principal axes of inertia. Their most probable configurations are very compact, while their extended configurations are essentially linear, as characterized by slowly diminishing distributions with the major axis of the distribution predominating. The large asymmetry of the chain configurations is seen in the ratios of first moments of three principal components of the gyration tensor: 12.41:3.11:1.00 for a linear chain with 29 beads and 7.12:2.70:1.00 for a circular chain with 49 segments, which are in agreement with those of Šolc and Stockmayer obtained from Monte Carlo simulations. The evaluations of the shape distribution functions for linear and circular macromolecules are based on two specifically designed algorithms, but a generic algorithm for the computation of the distribution functions for macromolecules of arbitrary complexity has also been developed. These algorithms make good use of the vector and parallel capabilities possessed by IBM 3090 systems. Asymptotic distributions, i.e., shape distribution functions for large values of the arguments, are given for flexible macromolecules of any structure such that the second smallest eigenvalue of the force constant matrix has odd degeneracy. The probability distributions of individual principal components and several of their linear combinations, including the squared radius of gyration, have also been studied. This work completes the description of asymmetric configurations of macromolecules in three-dimensional space. The integrals that have been evaluated in this work have applications beyond the physical sciences. Such integrals appear also in the diverse mathematical disciplines of multivariate statistics, harmonic analysis, and number theory.

Introduction

It is known that, in the absence of external forces, a flexible macromolecule may assume a very large number of configurations. Given an interaction potential for a macromolecular system, the probability distribution of measures of its shape may be calculated from classical statistical mechanics. The formulation of these measures is a purely mathematical problem.

Macromolecular systems have traditionally been modeled as collections of mass points, often called beads, and they have a definite structure ranging from a linear chain to an entire elastic network. If the long-range interactions among these beads are ignored, as in the case of a dilute polymer solution at the Θ temperature, a quadratic potential of mean force may be used to study their molecular configurations. This model, as simple as can be, has specifically been named the Gaussian model, since each configuration is equivalent to the trail left by a random walker with each step of the walk having a Gaussian probability.

The problem of determining shape distributions for Gaussian macromolecules had not been addressed until early 1971, when Šolc and Stockmayer[1] first published a short paper on the shape of a long linear chain. This work was continued in a separate paper by Šolc.[2] In their pioneering work, Šolc and Stockmayer were able to show a very strong departure of the average shape of the macromolecular chain from spherical symmetry, as had been earlier considered but not thoroughly investigated by Kuhn.[3] The explicit formulation of 3D shape distribution functions and their analytic solutions were first presented in 1977 by one of us.[4,5] In this work, it was proved that a Gaussian molecule is never spherically symmetric in a coordinate frame affixed to the principal axes of the molecule, and an analytic solution to the general shape distribution function was given as a series in zonal polynomials, which can be computed by following a certain algorithm. The mathematical context of this theory is found in multivariate statistics.[6]

It is important to note that the shape distribution function referred to here is, by definition, the joint probability density for finding a macromolecule in a particular configuration characterized by definite values of the three principal components of the gyration tensor of the molecule. It differs both from the segment density distributions that have been studied by Monte Carlo techniques[7] and from approximate forms of the joint probability distribution function, which are expressed as products of probability distribution functions for individual principal components.[8,9] Although these analytic forms recover the well-known Fixman result for the distribution function of the trace of the gyration tensor,[10] they are unable to yield a zero probability density for a configuration in which any two or all three principal components are equal.[4]

In this paper, we begin with a description of general three-dimensional shape distribution functions for Gaussian macromolecules,[5,11,12] giving first a generic algorithm that can be used to compute distribution functions for macromolecules of arbitrary complexity. Two specific computational algorithms for linear and circular macromolecules are then presented. Following this, we display results for numerically evaluated distribution functions for the latter types of macromolecules. Included here are several contracted distribution functions of individual principal components and their linear combinations, including the squared radius of gyration, for a circular chain with 11 beads. We obtain asymptotic distributions by use of the saddle-point method for macromolecules having an odd degeneracy of the second smallest eigenvalue of the force constant matrix. Finally, we discuss the numerical results with the aid of visual representations of the distribution functions for linear and circular macromolecules.

0024-9297/90/2223-4845$02.50/0

Computational Algorithms

The general three-dimensional shape distribution function for Gaussian macromolecules takes the form[4,5]

$$P(\tilde{S}_1,\tilde{S}_2,\tilde{S}_3) = (1/2\pi^2)\prod_{\alpha<\beta}^{3}|\tilde{S}_\alpha - \tilde{S}_\beta|\int\prod_{\alpha<\beta}^{3}|y_\alpha - y_\beta| \times \text{etr}\,(i\tilde{\mathbf{S}}_d\mathbf{h}\mathbf{y}\mathbf{h}')|\mathbf{1} + i\mathbf{y}\otimes\tilde{\mathbf{\Lambda}}^{-1}|^{-1/2}\,d\tilde{\mathbf{h}}\prod_{\alpha=1}^{3}dy_\alpha \quad (1)$$

where $\mathbf{y} = \text{diag}\,(y_1, y_2, y_3)$ with $-\infty \le y_3 \le y_2 \le y_1 \le \infty$, $\mathbf{h} \in SO(3)$ is the special orthogonal group acting on a three-dimensional space, and $d\tilde{\mathbf{h}}$ is the normalized Haar measure on SO(3). The diagonal matrix $\tilde{\mathbf{S}}_d$ consists of $\tilde{S}_1$, $\tilde{S}_2$, and $\tilde{S}_3$ with $\tilde{S}_j = 3S_j/2n^\alpha\langle l^2\rangle_0$, where n denotes the number of beads of the macromolecule, $\langle l^2\rangle_0$ is the mean-square unperturbed segmental length of the molecule, S_j is the jth principal component of the gyration tensor $\mathbf{S}$ for the molecule, with $0 \le S_3 \le S_2 \le S_1 \le \infty$, and α is a constant chosen such that the average trace of $\tilde{\mathbf{S}}_d$ has a finite, nonzero value in the large-n limit.[13] The matrix $\tilde{\mathbf{\Lambda}}$ is a diagonal matrix, the elements $\tilde{\lambda}_j (j = 1, 2, ..., n-1)$ of which are defined as $n^{\alpha+1}\lambda_j$, with the λ_j being the nonzero eigenvalues of the force constant matrix of the molecule. The shape distribution function $P(\tilde{S}_1,\tilde{S}_2,\tilde{S}_3)$ is related to the probability distribution for the gyration tensor through the equality $P(\mathbf{S})\,d\mathbf{S} = P(\tilde{S}_1,\tilde{S}_2,\tilde{S}_3)\,d\tilde{\mathbf{h}}\,d\tilde{S}_1\,d\tilde{S}_2\,d\tilde{S}_3$ and has been normalized, i.e.

$$\int_{0\le\tilde{S}_3\le\tilde{S}_2\le\tilde{S}_1\le\infty} P(\tilde{S}_1,\tilde{S}_2,\tilde{S}_3)\,d\tilde{S}_1\,d\tilde{S}_2\,d\tilde{S}_3 = 1 \quad (2)$$

To seek a generic algorithm for numerical evaluation of eq 1, we first note that, by a linear transformation of integration variables, eq 1 can be cast into the form[11]

$$P(\tilde{S}_1,\tilde{S}_2,\tilde{S}_3) = (3^5/2)|\tilde{\Lambda}|^{3/2}\prod_{\alpha<\beta}^{3}|\tilde{S}_\alpha - \tilde{S}_\beta| \times \int_{-\infty}^{\infty}dt\,e^{i\tilde{s}^2t}\int_0^\infty dx\,x\int_0^\infty dy\,y(x+y)\,W(x,y) \times D(t+2x+y)\,D(t-x+y)\,D(t-x-2y) \quad (3)$$

where $\tilde{s}^2 = \tilde{S}_1 + \tilde{S}_2 + \tilde{S}_3$, $|\tilde{\Lambda}| = \prod_{j=1}^{n-1}\tilde{\lambda}_j$ and $D(x) = \prod_{j=1}^{n-1}(\tilde{\lambda}_j + ix)^{-1/2}$. The function $W(x,y)$ is given by[11,14]

$$W(x,y) = \frac{1}{2\pi}\int_0^1 dr\int_0^{2\pi}d\alpha\,\exp[i3(x+y)\,\eta(r,\alpha)] \times J_0[3x\,\xi(r,\alpha)] \quad (4)$$

where $\eta(r,\alpha) = (\tilde{\Delta}_1/4)(1-r^2)\cos\alpha$ and $\tilde{\Delta}_1 = \tilde{S}_1 - \tilde{S}_2$. The function $\xi(r,\alpha)$ is defined as

$$\xi(r,\alpha) = (1/2)\{(\tilde{\Delta}_1 r\sin\alpha)^2 + [\tilde{\Delta}_2(1-r^2) + (\tilde{\Delta}_2/2) \times (1+r^2)\cos\alpha]^2\}^{1/2} \quad (5)$$

with $\tilde{\Delta}_2 = \tilde{S}_2 - \tilde{S}_3$, and $J_0(x)$ is the Bessel function of the first kind of zero order. Note that $W(x,y)$ may also be expressed as a series of zonal polynomials, i.e.

$$W(x,y) = \exp[-(i/3)\,\text{Tr}\,(\tilde{\mathbf{a}})\,\text{Tr}\,(\tilde{\mathbf{b}})]\;{}_0F_0^{(3)}(i\tilde{\mathbf{a}},\tilde{\mathbf{b}}) \quad (6)$$

where $\tilde{\mathbf{a}} = \text{diag}\,(\tilde{S}_1 - \tilde{S}_3,\ \tilde{S}_2 - \tilde{S}_3,\ 0)$, $\tilde{\mathbf{b}} = \text{diag}\,(3x + 3y,\ 3y,\ 0)$, and the two-matrix hypergeometric function ${}_0F_0^{(3)}(i\tilde{\mathbf{a}},\tilde{\mathbf{b}})$ is given by[6,15]

$${}_0F_0^{(3)}(i\tilde{\mathbf{a}},\tilde{\mathbf{b}}) = \sum_{m=0}^{\infty}\frac{i^m}{m!}\sum_{\mu}\frac{C_\mu(\tilde{\mathbf{a}})\,C_\mu(\tilde{\mathbf{b}})}{C_\mu(1_3)} \quad (7)$$

Here $C_\mu(\tilde{\mathbf{x}})$, with $\mathbf{x} = (\mathbf{a}, \mathbf{b})$, is the zonal polynomial and $\mu = (m_1, m_2)$ is a partition of m into no more than two parts such that $m_i \ge 0$ and $m_1 + m_2 = m$. Explicit expressions for $C_\mu(\tilde{\mathbf{x}})$ and $C_\mu(1_3)$ exist.[15] An effective way to compute ${}_0F_0^{(3)}(i\tilde{\mathbf{a}},\tilde{\mathbf{b}})$ based on the above series is to choose a new set of parameters such that all nonzero elements of the new diagonal matrix in the zonal polynomials are in the interval [0, 1].[16]

We now note that eq 3 can be rewritten as

$$P(\tilde{S}_1,\tilde{S}_2,\tilde{S}_3) = 3^5|\tilde{\Lambda}|^{3/2}\prod_{\alpha<\beta}^{3}|\tilde{S}_\alpha - \tilde{S}_\beta|\int_0^\infty dt\int_0^\infty dx\,x\int_0^\infty dy\,y \times (x+y)\,\text{Re}\,\{\exp(i\tilde{s}^2t)\,W(x,y)\,D(t+2x+y)\,D(t-x+y)\times D(t-x-2y)\} \quad (8)$$

which can be evaluated by numerical quadrature after a transformation of integration variables into spherical polar coordinates. However, a different approach leading also to only one final semiinfinite integral has proven to be much more effective, and this is as follows. Let us first define

$$V(t,x,y) = \exp(i\tilde{s}^2t)\,D(t+2x+y)\,D(t-x+y)\,D(t-x-2y) \quad (9)$$

and

$$U(t,x,y) = xy(x+y)\,W(x,y)[V(t,x,y) + \overline{V(t,y,x)}] \quad (10)$$

Then we can write eq 3 as

$$P(\tilde{S}_1,\tilde{S}_2,\tilde{S}_3) = (3^5/2)|\tilde{\Lambda}|^{3/2}\prod_{\alpha<\beta}^{3}|\tilde{S}_\alpha - \tilde{S}_\beta|\int_0^\infty dt\int_0^\infty dx\int_0^\infty dy\,U(t,x,y) \quad (11)$$

We now make the integral identity

$$\int_0^\infty dx\int_0^\infty dy\,U(t,x,y) = \int_0^\infty dx\,U(t,x,x) + \int_0^\infty dx\int_0^x dy\,U(t,x,y) + \int_0^\infty dy\int_0^y dx\,U(t,x,y) \quad (12)$$

which, along with the symmetry relations[11] $W(x,y) = \overline{W(y,x)}$ and $U(t,x,y) = \overline{U(t,y,x)}$, allows one to write

$$P(\tilde{S}_1,\tilde{S}_2,\tilde{S}_3) = (3^5/2)|\tilde{\Lambda}|^{3/2}\prod_{\alpha<\beta}^{3}|\tilde{S}_\alpha - \tilde{S}_\beta|\int_0^\infty dt\int_0^\infty dx\,x^3 \times F(t,x) \quad (13)$$

where $F(x,y)$ is defined as

$$F(x,y) = U(y,y) + 2y\,\text{Re}\int_0^1 dz\,U(x,y,yz) \quad (14)$$

Transforming the integration variables in eq 13 into plane polar coordinates, we finally arrive at

$$P(\tilde{S}_1,\tilde{S}_2,\tilde{S}_3) = (3^5/2)|\tilde{\Lambda}|^{3/2}\prod_{\alpha<\beta}^{3}|\tilde{S}_\alpha - \tilde{S}_\beta|\int_0^\infty d\rho\,\rho^4\int_0^1 dx \times (1-x^2)\,F(\rho x,\rho(1-x^2)^{1/2}) \quad (15)$$

which is the starting equation on which our generic algorithm is based.

The generic algorithm sought here consists of five major steps: (i) writing $D(x)$ as an exponential function of a sum (for a discrete eigenvalue spectrum) or an integral (for a continuous spectrum) and vector-adding the real and imaginary parts of the logarithm in the exponent in the case of a large number of discrete eigenvalues; (ii) using 10-point Gauss–Legendre quadrature to evaluate the integral over $z \in [0, 1]$ in eq 14; (iii) evaluating the double integral resulting from the W-function defined in eq 4 by use of the 10-point Fejér quadrature rule[17] and the IMSL

subroutine DQDAG;[17] (iv) performing 10-point Gaussian quadrature on the integration over $x \in [0, 1]$ in eq 15; (v) carrying out the final semiinfinite integration by first computing a series of data points for the integrand by using parallel processing and then performing either Romberg integration[18] with 5-point Lagrangian interpolation between these data points or Simpson integration[18] for this discrete set of values of the integrand. A Fortran code based on the above algorithm has been written and test run on IBM 3090 supercomputers.[19] The result from this algorithm was compared with that from other approaches for a linear chain with 11 beads, and good agreement between them was found.[19] We further note that with the use of the identity given in eq 12, approximate evaluation of eq 11 by a steepest descents technique may be done analogously to that used for evaluating two-dimensional shape distribution functions.[20]

For linear and circular chains, we have $\alpha = 1$ in the definitions of $\tilde{\mathbf{S}}_d$ and $\tilde{\mathbf{\Lambda}}$ given earlier,[13] and the eigenvalues of the force constant matrices are $\tilde{\lambda}_j = 4n^2 \sin^2 (\pi j/mn)$ where m takes the value 2 for linear chains and 1 for circular.[5] In these cases, we have found it advantageous to put eq 1 into the form[11]

$$P(\tilde{S}_1,\tilde{S}_2,\tilde{S}_3) = (1/16\pi^3)|\tilde{\Lambda}|^{3/2}\prod_{\alpha<\beta}^{3}|\tilde{S}_\alpha - \tilde{S}_\beta| \int_{-\infty}^{\infty}dt \int_0^{\infty}dx\, x \int_0^{\infty}dy\, y \int_{-1}^{1}dr \int_0^{2\pi}d\alpha \int_0^{2\pi}d\beta \times (x+y)\, F_1(t+x)\, F_2(t)\, F_3(t-y) \quad (16)$$

where $F_j(x) = D(x) \exp(iA_jx)$ with the A_j given by

$$A_1 = g(r) + a_+(1-r^2)\cos\beta + w(r,\beta)\cos\alpha \quad (17a)$$

$$A_2 = g(r) + a_-(1-r^2)\cos\beta - w(r,\beta)\cos\alpha \quad (17b)$$

and

$$A_3 = b_+ - 2ar^2 - b_-(1-r^2)\cos\alpha \quad (17c)$$

Here, $g(r) = (\tilde{s}^2 + \tilde{S}_3)/4 + ar^2$, $a_\pm = (\tilde{\Delta}_1 \pm 2\tilde{\Delta}_2)/4$, $b_\pm = (\tilde{S}_1 \pm \tilde{S}_2)/2$, and $w(r,\beta) = (\tilde{\Delta}_1/4)\{4r^2\sin^2\beta + [1 - r^2 + (1 + r^2)\cos\beta]^2\}^{1/2}$. As in the case of evaluating shape distribution functions for symmetric configurations of linear and circular macromolecules in three dimensions,[11] we may extend the integration limits of one of two semiinfinite integrals in eq 16 to $[-\infty, \infty]$. This permits one to use the residue theorem twice to contour integrate the resulting two infinite integrals for both types of molecules with n greater than 5. For linear macromolecules, vertical branch cuts may be used in the contour integration.[11,19,21] In the case of circular macromolecules with an odd number of beads, only simple poles are present in the integrands. When n is even for circular macromolecules, however, additional integrations along proper branch cuts need to be done.[19] Here, we limit ourselves to the odd-n cases for both linear and circular macromolecules.

After carrying out consecutively two contour integrations for circular macromolecules, we are left with

$$P(\tilde{S}_1,\tilde{S}_2,\tilde{S}_3) = (2/\pi)|\tilde{\Lambda}|^{1/2}\prod_{\alpha<\beta}^{3}|\tilde{S}_\alpha - \tilde{S}_\beta| \int_0^{\infty}dx \int_0^{1}dr \int_0^{2\pi}d\alpha\, J_0(\xi x) \sum_{j\neq k}^{(n-1)/2} B_jB_k(\tilde{\lambda}_k - \tilde{\lambda}_j) \times \exp[(2\eta - \tilde{s}^2)\tilde{\lambda}_j - 2\eta\tilde{\lambda}_k]\ \mathrm{Im}\ \frac{e^{i\eta x}}{\prod_{l\neq j,k}^{n-1}(\tilde{\lambda}_l - \tilde{\lambda}_k + ix)} \quad (18)$$

where $\eta(r,\alpha)$ and $\xi(r,\alpha)$ are the same as defined before, and B_k is given by

$$B_k = \tilde{\lambda}_k / \prod_{l\neq k}^{(n-1)/2}(1 - \tilde{\lambda}_k/\tilde{\lambda}_l) \quad (19)$$

For linear macromolecules, we have

$$P(\tilde{S}_1,\tilde{S}_2,\tilde{S}_3) = (8/\pi^3)|\tilde{\Lambda}|^{1/2}\prod_{\alpha<\beta}^{3}|\tilde{S}_\alpha - \tilde{S}_\beta| \int_0^{\infty}dx \int_0^{1}dr \int_0^{2\pi}d\alpha\, J_0(\xi x) \times \sum_{j\neq k}^{(n-1)/2}(-1)^{j+k}\int_0^{\pi/2}d\theta \int_0^{\pi/2}d\varphi\, \beta(\epsilon_k - \epsilon_j)D_jD_k \times \exp[(2\eta - \tilde{s}^2)\epsilon_j - 2\eta\epsilon_k]\ \mathrm{Re}\ \frac{x(\epsilon_j - \epsilon_k + ix)}{\prod_{l=1}^{n-2}(\tilde{\lambda}_l - \epsilon_k + ix)^{1/2}}e^{i\eta x} \quad (20)$$

where $\epsilon_k = \tilde{\lambda}_{2k-1} + (\tilde{\lambda}_{2k} - \tilde{\lambda}_{2k-1})\sin^2\varphi$ and D_k is defined as

$$D_k = (\tilde{\lambda}_{2k}\tilde{\lambda}_{2k-1})^{1/2}\prod_{l\neq 2k,2k-1}^{n-1}|\epsilon_k/\tilde{\lambda}_l - 1|^{-1/2} \quad (21)$$

Equations 18 and 20 may be evaluated analogously to that used for eq 15. However, more can be done on both equations, as one may use the exponential integral functions to represent the semiinfinite integrals in eq 18 and reduce the improper integrals in eq 20 to a sum of simple integrals.[11,19] We have done this, and the resulting equations are to be presented in next section. Here, to emphasize algorithm development, we merely indicate that the final equations for both linear and circular macromolecules can be put into the form[19]

$$P(\tilde{S}_1,\tilde{S}_2,\tilde{S}_3) = \int_0^{1}dx \int_0^{1}dy \sum_{j=1}^{3} W_j(x,y)\, G_j(x,y) \quad (22)$$

where W_j and G_j are functions of $\tilde{S}_1$, $\tilde{S}_2$, and $\tilde{S}_3$. We now notice that since W_j and G_j are independent of each other for a given set of values of $\tilde{S}_1$, $\tilde{S}_2$, and $\tilde{S}_3$, they may be computed simultaneously across six processors of IBM 3090 supercomputers. As before, vector addition may be used in certain partial-sum operations. We have written and test run four different versions of Fortran codes involving four different combinations among the four available modes: scalar, serial, vector, and parallel.

Evaluations and Results

Evaluation of multiple integrals with oscillatory integrands remains a great challenge. This fact was learned while searching for both stable and efficient algorithms for evaluating three-dimensional shape distribution functions. Attempts at direct evaluations of the improper integrals contained in these functions by means of the IMSL subroutine DQAWF[17,22] proved unsuccessful with respect to both the accuracy of the results and the computing time consumed. The computational algorithm presented in the previous section avoids, to a certain degree, oscillation of the integrands, thereby ensuring the stability of the algorithms and subsequently the accuracy of the computed results. It is imperative to avoid, whenever possible, rapidly oscillating integrands to ensure reasonable accuracy and efficiency. In this respect, we have found that it is advantageous to further analytically reduce the integrals in eqs 18 and 20 to more slowly oscillating or even monotonic functions.

For circular macromolecules, eq 18 is transformed into

the form

$$P(\tilde{S}_1,\tilde{S}_2,\tilde{S}_3) = (1/\pi^2)\prod_{\alpha<\beta}^{3}|\tilde{S}_\alpha - \tilde{S}_\beta| \times \int_0^1 dr \int_0^{2\pi} d\alpha \int_0^{2\pi} d\beta \sum_{j\neq k\neq l}^{(n-1)/2} B_j B_k B_l (\tilde{\lambda}_j - \tilde{\lambda}_k)(\tilde{\lambda}_j - \tilde{\lambda}_l) \times (\tilde{\lambda}_l - \tilde{\lambda}_k) \exp[(2\eta - \tilde{s}^2)\tilde{\lambda}_j - 2\eta\tilde{\lambda}_k] E[(\tilde{\lambda}_l - \tilde{\lambda}_k)(\eta + \xi\cos\beta)] \quad (23)$$

where $E(x) = \exp(-x)\,\mathrm{Ei}(x)$ and $\mathrm{Ei}(x)$ is the exponential integral function,[23] which can be evaluated either by using the IMSL subroutine MMDEI[17] or by writing one's own codes[19] based on polynomial approximations to $E_1(x)$[23,24] with 10-point Gaussian quadrature for the evaluation of $\mathrm{Shi}(x)$.[24] For linear macromolecules, eq 20 gives

$$P(\tilde{S}_1,\tilde{S}_2,\tilde{S}_3) = (16/\pi^3)\prod_{\alpha<\beta}^{3}|\tilde{S}_\alpha - \tilde{S}_\beta| \int_0^1 dr \int_0^{2\pi} d\alpha \times \sum_{j\neq k}^{(n-1)/2} (-1)^{j+k} \int_0^{\pi/2} d\theta \int_0^{\pi/2} d\varphi\, (\epsilon_k - \epsilon_j) D_j D_k \times \exp[(2\eta - \tilde{s}^2)\epsilon_j - 2\eta\epsilon_k] \int_0^{\pi/2} d\beta\, \{(-1)^{(n-1)/2}(\sigma_k + \epsilon_k - \epsilon_j) \times \sec^2\beta H(\sigma_k,\xi,\eta)\, D(\tilde{\lambda}_{n-1} - \epsilon_k, \sigma_k + \epsilon_k) + \sum_{l=k}^{(n-3)/2} (-1)^l (\epsilon_{l+1/2} - \epsilon_j) D_{l+1/2} H(\epsilon_{l+1/2} - \epsilon_k, \xi, \eta)\} \quad (24)$$

where $\sigma_k = (\tilde{\lambda}_{n-1} - \epsilon_k)\sec^2\beta$, and $D(a,b)$ and $H(x,a,b)$ are defined as

$$D(a,b) = (\tilde{\lambda}_{n-1}a)^{1/2}\prod_{l=1}^{n-2}(b/\tilde{\lambda}_l - 1)^{-1/2} \quad (25)$$

and

$$H(x,a,b) = xI_0(ax)\exp(-bx) \quad (26)$$

Here, $I_0(x)$ is the modified Bessel function of the first kind of zero order. The function $H(x,a,b)$ may be calculated with use of the IMSL subroutine DBSIOE[17] or by means of polynomial approximations to the exponentially scaled $I_0(x)$.[19,23-25]

We now make a crucial observation that both eq 23 and eq 24 can be significantly simplified by writing the shape distribution function as a double integral with the integrand consisting of six independent components, each involving a single or double sum of simple integrals. The resulting equations take the form given in eq 22. By doing so, we have achieved an efficiency several orders of magnitude greater. Still more can be done, however, as one can transform all angular integrals in eqs 23 and 24 into integrals over the intervals [0, 1]. These integrals may then be evaluated by use of simple quadrature rules. We have found that 10-point Gaussian or Fejér quadrature works very efficiently with sufficient accuracy. We have therefore achieved the goal of evaluating the distribution functions by 10-point quadrature, which requires the least computing time. In addition, this procedure results in a relatively uniform distribution of CPU time required for various values of the parameters, which cannot be achieved by another method, such as Romberg integration.

With stable and efficient algorithms for calculating $P(\tilde{S}_1,\tilde{S}_2,\tilde{S}_3)$ for linear and circular macromolecules over the entire range of variables at hand, we may then proceed to evaluate distribution functions of quantities that are arbitrary functions of the three principal components. One such quantity, the radius of gyration, $\tilde{s}^2 = \tilde{S}_1 + \tilde{S}_2 + \tilde{S}_3$, is of particular interest since its probability distribution has been separately studied.[4,10,13] Other quantities that may be worthy of consideration are the individual principal components and those that are the difference between the sum of any two principal components and twice the remaining one. The distribution functions of all these quantities share the generic form

$$P(\sigma)\,d\sigma = d\sigma \int_{0\le\tilde{S}_3\le\tilde{S}_2\le\tilde{S}_1\le\infty} \delta(\alpha\tilde{S}_1 + \beta\tilde{S}_2 + \gamma\tilde{S}_3 - \sigma) \times P(\tilde{S}_1,\tilde{S}_2,\tilde{S}_3)\,d\tilde{S}_1\,d\tilde{S}_2\,d\tilde{S}_3 \quad (27)$$

where α, β, and γ are any real numbers and $\delta(x)$ is the delta function. Equation 27 may then be transformed into

$$P(\sigma)\,d\sigma = (d\sigma/C)\int P[\tilde{S}_1(\sigma,x,y),\tilde{S}_2(\sigma,x,y),\tilde{S}_3(\sigma,x,y)]\,dx\,dy \quad (28)$$

through various linear transformations, with C denoting the Jacobian of each transformation. One such transformation, i.e.

$$\begin{cases} x = \alpha\tilde{S}_1 - \beta\tilde{S}_2 \\ y = \beta\tilde{S} - \gamma\tilde{S}_3 \\ \sigma = \alpha\tilde{S}_1 + \beta\tilde{S}_2 + \gamma\tilde{S}_3 \end{cases} \quad (29)$$

is applicable for any nonzero α, β, and γ. The domain of integration in eq 28 is either $a_1(\sigma) \le x \le b_1(\sigma)$ and $a_2(\sigma,x) \le y \le b_2(\sigma,x)$ or $a_3(\sigma) \le y \le b_3(\sigma)$ and $a_4(\sigma,y) \le x \le b_4(\sigma,y)$, where the a_j and b_j are determined by the original integration ranges $0 \le \tilde{S}_3 \le \tilde{S}_2 \le \tilde{S}_1 \le \infty$. Taking the case $\sigma = \tilde{S}_1 + \tilde{S}_2 + \tilde{S}_3 = \tilde{s}^2$, for example, we have, from eqs 28 and 29

$$P(\tilde{s}^2)\,d\tilde{s}^2 = \frac{d\tilde{s}^2}{3}\int_0^{\tilde{s}^2} dx \int_0^{(\tilde{s}^2-x)/2} dy \times P\left(\frac{\tilde{s}^2+2x+y}{3}, \frac{\tilde{s}^2-x+y}{3}, \frac{\tilde{s}^2-x-2y}{3}\right) \quad (30)$$

The expressions for a_j and b_j for seven selected sets of values of α, β, and γ are summarized in Table I.

Table I
Definitions of a_m and b_m for Selected Linear Combinations of $\tilde{S}_1$, $\tilde{S}_2$, and $\tilde{S}_3$

α	β	γ	m	a_m	b_m	a_{m+1}	b_{m+1}
1	0	0	1	0	σ	0	$\sigma - x$
0	1	0	1	0	∞	0	σ
0	0	1	1	0	∞	0	∞
1	1	1	1	0	σ	0	$(\sigma - x)/2$
2	−1	−1	3	$-\sigma$	0	$\sigma - 2y$	∞
1	−2	1	3	$-\infty$	0	$\sigma - y$	$\sigma - 2y$
			3[a]	$-\infty$	2σ	$-y - \sigma$	$\sigma - 2y$
1	1	−2	1	0	σ	$(\sigma - x)/2$	∞

[a] This row of entries is for $\tilde{S}_1 - 2\tilde{S}_2 + \tilde{S}_3 \le 0$.

Computations of shape distribution functions based on eq 23 for circular macromolecules with n = 11, 49, and 61 and eq 24 for linear macromolecules with n = 11, 29, and 61 were done on IBM 3090 supercomputers. The results are visualized in Figures 1–11. In Tables II–V are numerical values associated with the most probable (symbolized by an asterisk) and the average (denoted by $\langle\ \rangle$) configurations of the molecules. Figure 12 illustrates the distribution function of the radius of gyration for a circular macromolecule with 11 beads calculated according to eq 30, which is also compared with that from a previous explicit evaluation.[26] Numerical results from sample calculations of the distribution functions of individual principal components and three of their linear combinations according to eq 28 are plotted in Figures 13 and 14 for an 11-bead circular macromolecule. The most

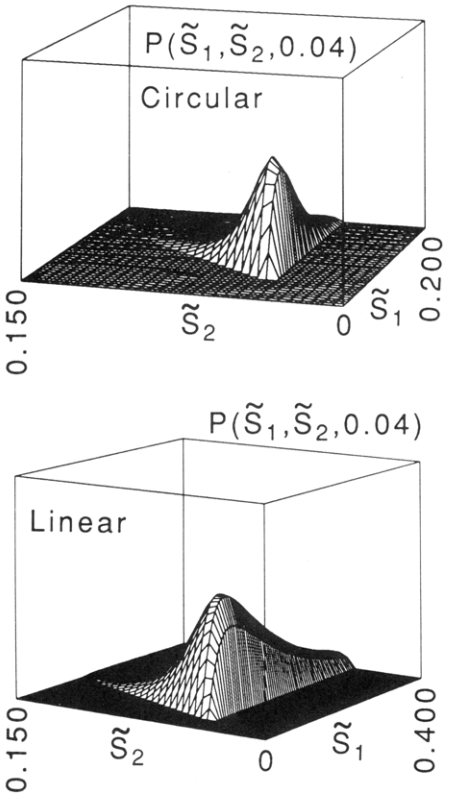

Figure 1. Shape distribution function $P(\tilde{S}_1,\tilde{S}_2,\tilde{S}_3)$ at $\tilde{S}_3 = 0.04$ for 11-bead circular (top) and linear (bottom) macromolecules, where $\tilde{S}_j = 3S_j/2n\langle l^2\rangle_0$. S_j is the jth principal component and $\langle l^2\rangle_0$ is the mean-square unperturbed length of the segment connecting neighboring beads.

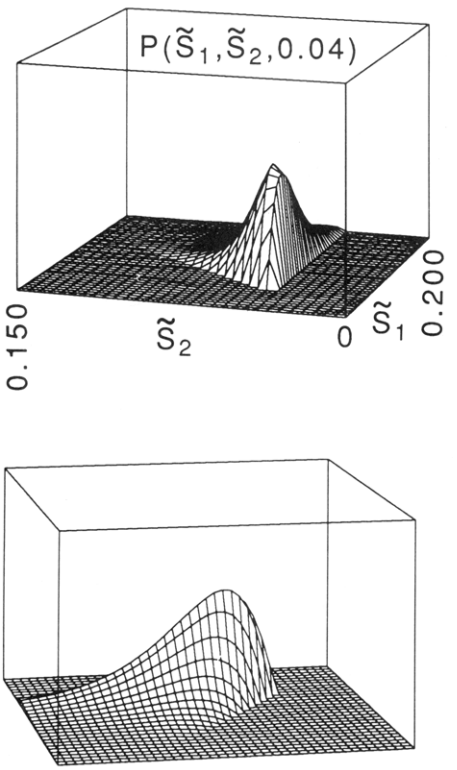

Figure 2. (Top) Shape distribution function with $\tilde{S}_3$ fixed at 0.04 for circular macromolecules with $n = 61$. (Bottom) The same distribution function viewed from a different angle.

probable configurations and the first and second moments of these distributions are tabulated in Table VI.

Much more CPU time is required for the evaluation of the distribution function for linear macromolecules than for circular ones of the same n. For a given type of structure, evaluation of the distribution function for large n consumes more computing time than for small n. Vectorization of scalar codes for linear macromolecules reduces CPU time only for large n. Parallelization of serial codes is realized on both task and case levels. The total CPU time on IBM 3090 supercomputers used in this work exceeded 1 month.

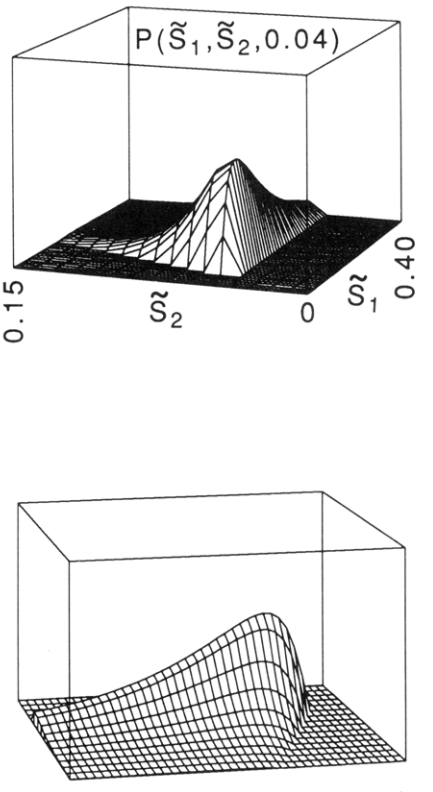

Figure 3. (Top) Shape distribution function at $\tilde{S}_3 = 0.04$ for linear macromolecules with $n = 61$. (Bottom) The same distribution function viewed from another angle.

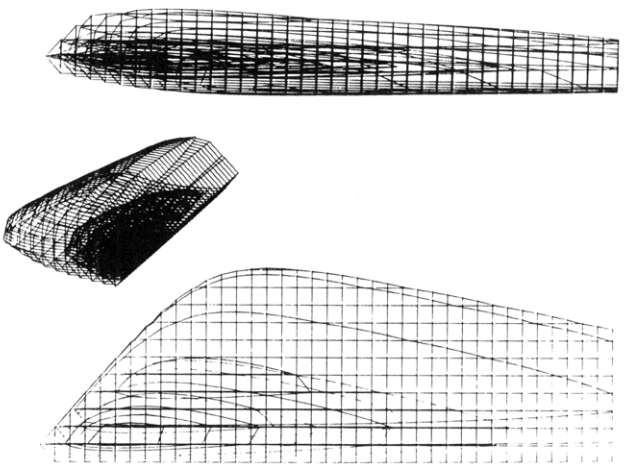

Figure 4. Shape distribution function $P(\tilde{S}_1,\tilde{S}_2,\tilde{S}_3)$ for 29-bead linear macromolecules over the entire variable domain, with $\tilde{S}_1$, $\tilde{S}_2$, and $\tilde{S}_3$ running from 0 to 0.40, 0.15, and 0.10, respectively, with a grid size of 0.01. The 4D data are visualized as equal probability density contour plots viewed from three different angles. Notice the differences among the three figures in depicting the region of the highest probability density.

Asymptotic Distributions

The force constant matrices for finite Gaussian macromolecules have discrete eigenvalue spectra. One can find asymptotic expansions of their distribution functions for large values of the three principal components. In fact,

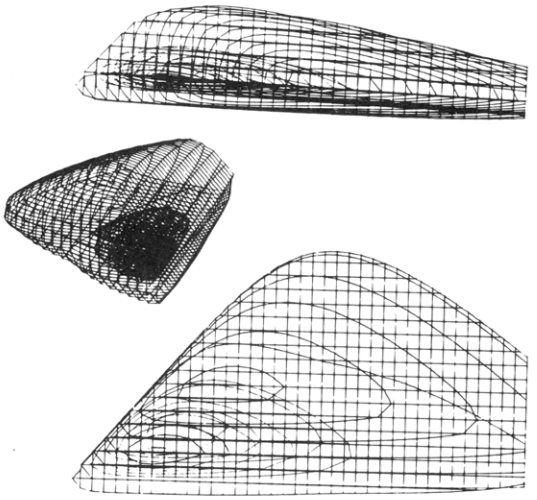

Figure 5. Shape distribution function for circular macromolecules with $n = 49$ over the entire variable domain. Otherwise the same as Figure 4 except that $\tilde{S}_1$ and $\tilde{S}_3$ run from 0 to 0.200 and 0.050, respectively, and that the grid size is 0.005 for all the three reduced principal components. Notice the differences in the overall shapes between these plots and those in Figure 4.

the leading terms of such expansions for many of these distribution functions, when formulated as integrals over the orthogonal group, can be found in the mathematical literature.[6] Here we extend our previous treatment for obtaining asymptotic distributions for certain types of macromolecules in two-dimensional space[20] to three-dimensional shape distribution functions.

We first write eq 1, for convenience, as

$$P(\tilde{S}_1,\tilde{S}_2,\tilde{S}_3) = (1/2\pi^2)\prod_{\alpha<\beta}^{3}|\tilde{S}_\alpha - \tilde{S}_\beta| \int_{-\infty\le y_3\le y_2\le y_1\le\infty} F(\tilde{\mathbf{S}}_d,\mathbf{y})\prod_{\alpha=1}^{3}dy_\alpha \quad (31)$$

where $F(\tilde{\mathbf{S}}_d,\mathbf{y})$ is given by

$$F(\tilde{\mathbf{S}}_d,\mathbf{y}) = \prod_{\alpha<\beta}^{3}|y_\alpha - y_\beta||1 + i\mathbf{y}\otimes\tilde{\Lambda}^{-1}|^{-1/2} \times \int_{SO(3)}\mathrm{etr}\,(i\tilde{\mathbf{S}}_d\mathbf{h}\mathbf{y}\mathbf{h}')\,d\tilde{\mathbf{h}} \quad (32)$$

Now let $\mathbf{y}_s$ denote the saddle point $(y_{1,s}, y_{2,s}, y_{3,s})$, if one exists. Application of the multidimensional saddle-point method[6] to eq 31 then gives

$$P(\tilde{S}_1,\tilde{S}_2,\tilde{S}_3) \simeq (2/\pi)^{1/2}\prod_{\alpha<\beta}^{3}|\tilde{S}_\alpha - \tilde{S}_\beta|\,F(\tilde{\mathbf{S}}_d,\mathbf{y}_s)[H(\mathbf{y}_s)]^{-1/2} \quad (33)$$

where $\mathbf{y}_s$ satisfies the system of equations

$$\mathbf{f}(\mathbf{y}_s) = 0 \quad (34)$$

with the constraints

$$F(\tilde{\mathbf{S}}_d,\mathbf{y}_s) > 0 \quad (35)$$

and

$$H(\mathbf{y}_s) = -\det\begin{pmatrix} f_{11}(\mathbf{y}_s) & f_{12}(\mathbf{y}_s) & f_{13}(\mathbf{y}_s) \\ f_{21}(\mathbf{y}_s) & f_{22}(\mathbf{y}_s) & f_{23}(\mathbf{y}_s) \\ f_{31}(\mathbf{y}_s) & f_{32}(\mathbf{y}_s) & f_{33}(\mathbf{y}_s) \end{pmatrix} > 0 \quad (36)$$

Here, (f_1, f_2, f_3) is the transpose of $\mathbf{f}$ with $f_i = \partial \ln F/\partial y_i$, and $f_{ij} = \partial^2 \ln F/\partial y_i \partial y_j$.

Just as for asymptotic distributions in two dimensions, we find that when the second smallest eigenvalue of the force constant matrix for a macromolecule has odd degeneracy, the asymptotic regions of three-dimensional shape distribution functions correspond to the saddle points $\mathbf{x}_s = (\kappa_1 - \epsilon_1, \kappa_2 - \epsilon_2, \kappa_3 - \epsilon_3) = -i\mathbf{y}_s$, where $\epsilon_j = \omega_j/2\tilde{S}_j$ and ω_j is the degeneracy of the jth distinct eigenvalue κ_j ($j = 1, 2, ..., p$) of $\tilde{\Lambda}$, ordered such that $\kappa_j < \kappa_{j+1}$. Then, at $\mathbf{y}_s = i\mathbf{x}_s$ for large $\tilde{S}_j$, we have

$$H(\mathbf{y}_s) \simeq \prod_{j=1}^{3}\frac{\omega_j}{2\epsilon_j^2} \quad (37)$$

and

$$F(\tilde{\mathbf{S}}_d,\mathbf{y}_s) \simeq (1/2\sqrt{\pi}C_1)\prod_{\alpha<\beta}^{3}|\tilde{S}_\alpha - \tilde{S}_\beta|^{-1/2}\prod_{j=1}^{3}\epsilon_j^{-\omega_j/2}\exp(-\kappa_j\tilde{S}_j) \quad (38)$$

where C_1 is given by

$$C_1 = (\kappa_2/\kappa_1 - 1)^{\omega_2/2}(\kappa_3/\kappa_2 - 1)^{\omega_3/2}(\kappa_3/\kappa_1 - 1)^{\omega_3/2}\prod_{\alpha<\beta}^{3}|\kappa_\alpha - \kappa_\beta|^{-1/2}\prod_{\alpha=1}^{3}\prod_{j=\alpha+1}^{p}(1 - \kappa_\alpha/\kappa_j)^{\omega_\alpha} \quad (39)$$

In eq 38, the asymptotic form of the integral over SO(3) given in ref 6 as Theorem 9.5.2 has been used. Substitution of eqs 37 and 38 into eq 33 then yields

$$P(\tilde{S}_1,\tilde{S}_2,\tilde{S}_3) \sim (2/\pi)C_2(1 - \tilde{S}_2/\tilde{S}_1)^{1/2}(1 - \tilde{S}_3/\tilde{S}_2)^{1/2} \times (\tilde{S}_1/\tilde{S}_3 - 1)^{1/2}\prod_{\alpha=1}^{3}\tilde{S}_\alpha^{(\omega_\alpha-1)/2}\exp(-\kappa_\alpha\tilde{S}_\alpha) \quad (40)$$

where $C_2 = C_1^{-1}(e/2)^{3/2}\prod_{j=1}^{3}\omega_j^{(\omega_j-1)/2}$. We note that, apart from a constant factor, the asymptotic distribution functions as given by eq 40 for any odd ω_2 reduce to the earlier result[5] for linear macromolecules.

Discussion and Conclusion

In the coordinate frame of the principal axes of inertia, here called the principal coordinate frame, there exist only three types of configurations—spheres, ellipsoids of revolution, and asymmetric ellipsoids. Previous studies[4,5,11] have shown that when double or triple degeneracy occurs in the three principal components of the gyration tensor for a macromolecule, the probability distributions are different from those in the nondegenerate case. The former corresponds to probability distributions for symmetric configurations, while the latter, which is studied in this paper, is defined only for asymmetric ones; i.e., the probability density is zero when any two or all three principal components become identical and takes a finite, nonzero value when the three principal components are distinct. This is seen in eq 1. During the past several years, it has been learned that this fundamental feature of shape distribution functions for Gaussian molecules can only be captured by explicitly formulating the shape distribution function for a given type of configuration. Any study that avoids directly dealing with the shape distribution functions themselves will therefore be unable to reveal this very important feature of the distributions.

Once the asymmetric nature of shape distributions is known, what is left is to characterize this asymmetry by average values of properly chosen mathematical objects (see, for example, ref 8) or to study directly the distribution functions themselves. The latter is certainly the most desired yet most formidable task, and that is what we have undertaken as described in the previous sections, the results of which we will discuss shortly.

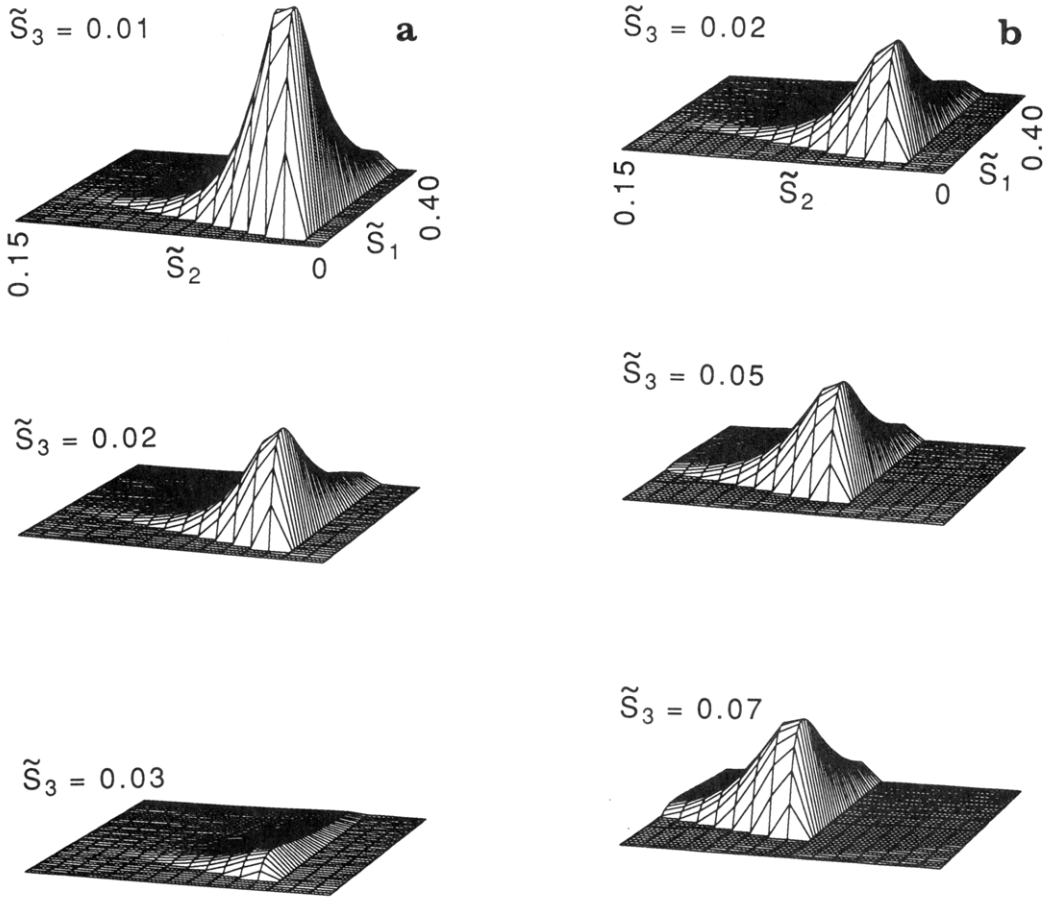

Figure 6. (a) Shape distribution function for 29-bead linear macromolecules, which evolves as $\tilde{S}_3$ changes from 0.01 to 0.03 with a step size of 0.01. (b) The same function for three fixed values of $\tilde{S}_3$, all plots unscaled in height. Note the changes in location of each maximum as $\tilde{S}_3$ varies.

As a first-order measure of asymmetry, the ratios of first moments of principal components, $\langle S_1\rangle:\langle S_2\rangle:\langle S_3\rangle$ for linear and circular macromolecules of two different lengths, see Table I, immediately indicate that the distributions for these molecules are highly asymmetric, with those for circular macromolecules considerably less so, as expected. The large ratios of $\langle S_1\rangle:\langle S_2\rangle$ and $\langle S_1\rangle:\langle S_3\rangle$ for both types of macromolecules show that their shape distribution functions have long tails along the largest principal component axis. These numerical values compare quite well with those obtained by Monte Carlo methods for long linear and circular macromolecules; i.e., $\langle S_1\rangle:\langle S_2\rangle:\langle S_3\rangle$ = 11.7:2.7:1 for a 51-bead linear chain[1,7] and 6.5:2.35:1 for a 50-bead circular chain.[27] The most probable configurations for these macromolecules are found to be very compact with respect to the mean-square dimensions, with $^*\tilde{s}^2$ being 62.5% and 44.6% of $\langle\tilde{s}^2\rangle_0$ for linear macromolecules with n = 29 and circular chains with n = 49, respectively.

The large asymmetry, compactness of the most probable configurations, and long tails of the distributions along the major axis as implied by the average and most probable principal components obtained from the complete shape distribution functions may be more readily observed from the graphical displays of these functions themselves. Figure 1 illustrates the distribution functions for both linear and circular macromolecules with the smallest reduced principal component fixed at 0.04. We notice that these distributions resemble those in two dimensions.[28] This large asymmetry was also seen in the distributions for ellipsoids of revolution studied earlier by us.[11] The retention of the features of low-dimensional distributions in higher dimensional ones seems to be a rather general property of shape distributions for Gaussian macromolecules. Figures 2 and 3 display the distribution functions at n = 61 for circular and linear macromolecules, respectively. As is seen from these figures and Table V, the distribution functions have a very weak dependence on the size of the macromolecule when n exceeds 11.

Figures 4 and 5 are four-dimensional contour plots of the probability densities $P(\tilde{S}_1,\tilde{S}_2,\tilde{S}_3)$ for both linear and circular macromolecules. They visualize the 4D data as equal probability density contour plots, viewed from selected angles in 3D space. These figures are to be understood in the same way that one interprets electron density maps, $\Psi^2(x,y,z)$, for atoms and molecules. It is evident from these pictures that Gaussian molecules are most likely found within a small region of the variable domains that is not far from the origin and that they are highly anisometric when viewed from the principal coordinate frame. Circular macromolecules are somewhat less asymmetric than are linear ones. Moreover, linear macromolecules are seen to be more highly extended than circular ones.

Figures 6–8 show the evolution of the shape distribution functions for linear macromolecules with n = 29 as one of three reduced principal components increases. It is seen that the peak diminishes most rapidly along the smallest principal component axis. Less rapid diminishing of the peak is found when $\tilde{S}_2$ varies, as seen in Figure 7. Figure 8 shows that the distribution function varies most slowly with the largest reduced principal component. These highly extended configurations are apt to be of great importance ot the translational motion of chains, as studied for example in the electrophoresis of DNA.[29] Figures 9–11 illustrate how the shape distribution function for a 49-

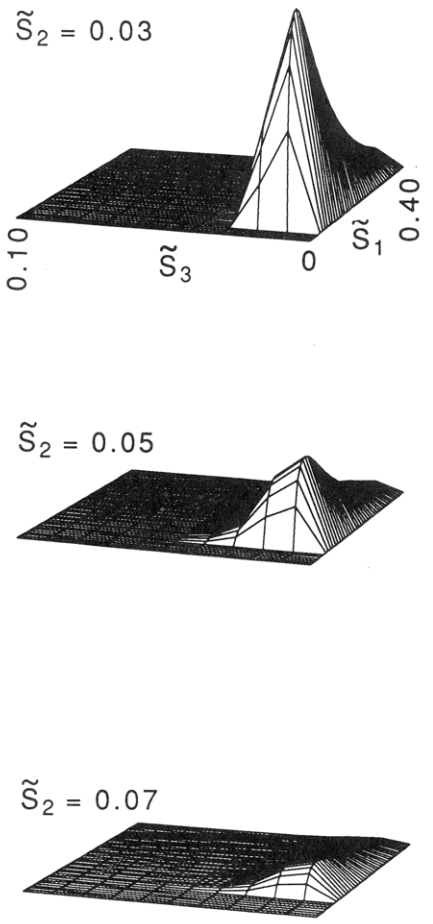

Figure 7. Shape distribution function for linear macromolecules with $n = 29$, which evolves as $\tilde{S}_2$ increases from 0.03 to 0.07 with step size 0.02.

bead circular macromolecule evolves as one of the three principal components increases. We notice that the peak of the distribution function in each case diminishes much more rapidly than that for linear macromolecules. The locations of all the peaks seen in Figures 6–11 can be found in Tables III–V.

The shape distribution functions for large values of the reduced principal components for the case of odd degeneracy of the second smallest eigenvalue of the force constant matrix of Gaussian macromolecules studied here have important applications in the theory of elasticity,[12,30] and they provide simple asymptotic expressions for linear and certain star-shaped macromolecules. Furthermore, since these asymptotic forms are exact, they can be used to test approximations to shape distribution functions. They may also be used to extend the domains beyond those that are reasonably integratable by numerical methods. As a final note, it should be noted that, even in the asymptotic region, shape distribution functions are not separable into distribution functions of individual principal components.

A study of the probability distributions of certain linear combinations of three principal components not only serves to check the accuracy of the numerical work but it provides us with some specific statistical information concerning these quantities. Figure 12 shows the distribution function of the squared radius of gyration for circular macromolecules with $n = 11$ as calculated both from eq 30 and from eq 10c of ref 26. Since the two curves are indistinguishable in the figure, both are to be regarded as identical to within numerical accuracy. It is now clear that the distribution function of the squared radius of gyration, the trace of the gyration tensor, is equivalent to that of

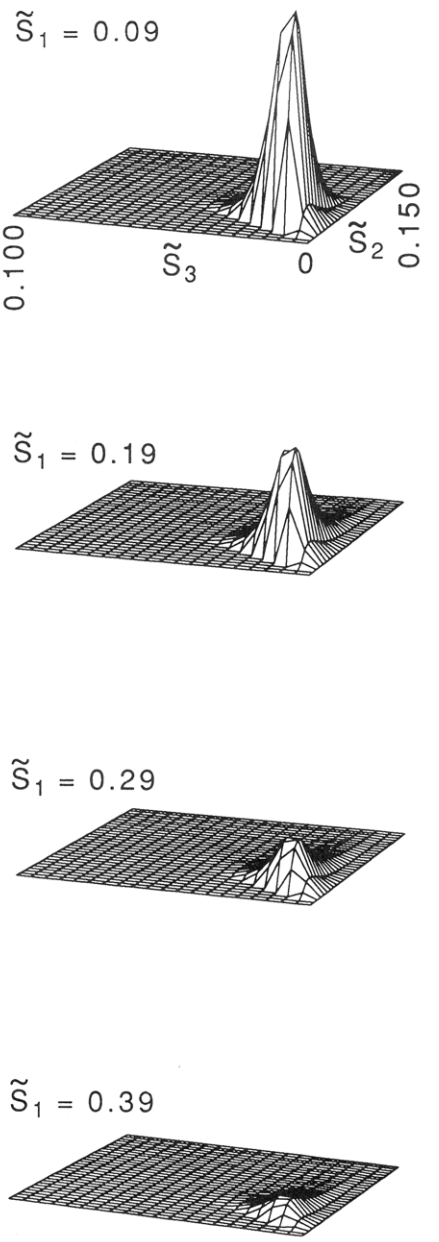

Figure 8. Shape distribution function for 29-bead linear macromolecules, which evolves as $\tilde{S}_1$ varies from 0.09 to 0.39 with a step size of 0.10.

the sum of the three principal components only for asymmetric configurations. This is not the case for ellipsoids of revolution.[11] The reason for this is that distributions on the boundary of the domain $S_1 > S_2 > S_3$ are not to be compared directly with those within the domain.

From Figure 13, it is seen that for circular macromolecules with $n = 11$ the distribution for the smallest principal component is quite peaked, while that for the largest is broad. The most probable and average values of $\tilde{S}_1$, the largest reduced principal component, are slightly less than those predicted by the approximate form given in ref 8, which is in agreement with Monte Carlo simulations.[9] It must be noted here that $P(\tilde{S}_1)$ for linear chains in three dimensions may be well approximated by the form $\tilde{S}_1^{1/2} \exp(-\kappa_1 \tilde{S}_1)$, but it is certainly incorrect to use the general expression $\tilde{S}_i^{1/2} \exp(-\kappa_i \tilde{S}_i)$ to approximate the distribution functions of the smaller principal components, even in the large-n limit, as is clear from close inspection of the results presented in Table II.

In Figure 14 three distribution functions for simple linear combinations of the three principal components are plotted

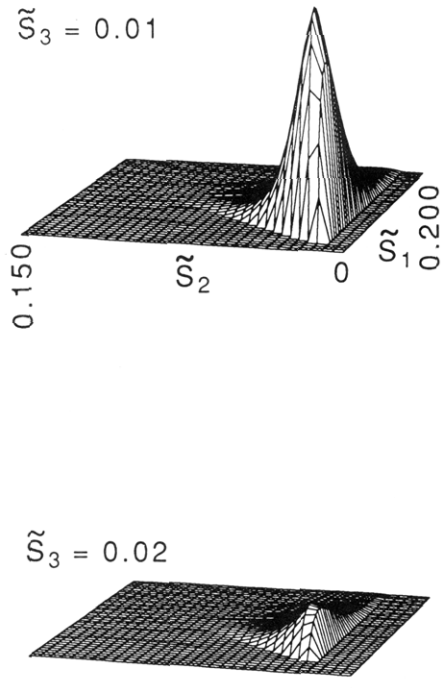

Figure 9. Shape distribution function for circular macromolecules with $n = 49$, which evolves as $\tilde{S}_3$ increases from 0.01 to 0.02.

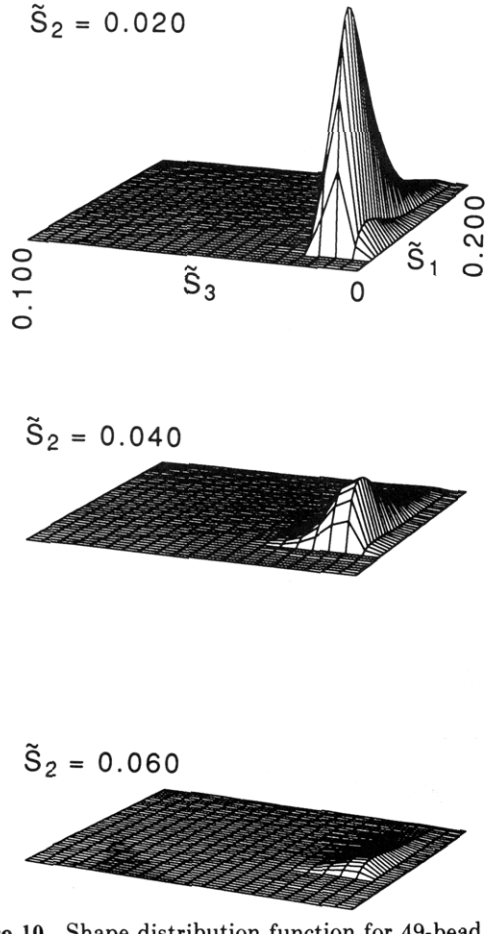

Figure 10. Shape distribution function for 49-bead circular macromolecules, which evolves as $\tilde{S}_2$ varies from 0.020 to 0.060 with a step size of 0.020.

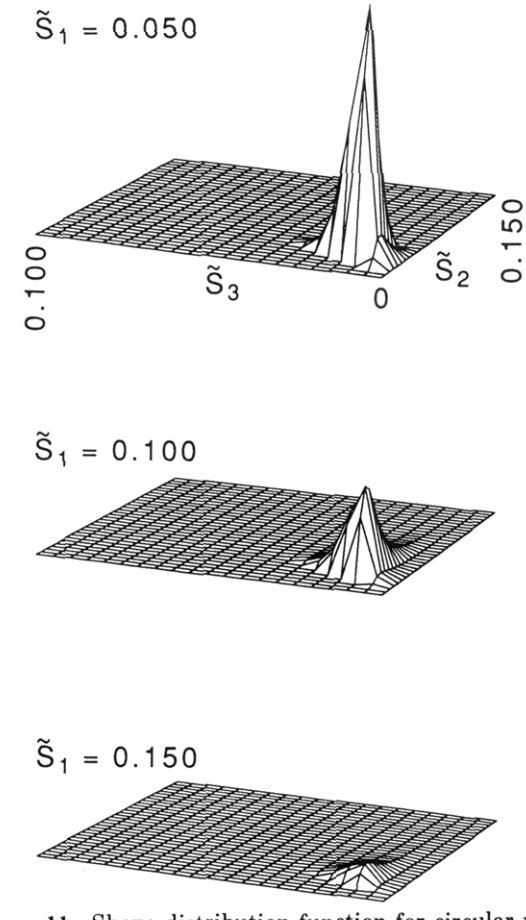

Figure 11. Shape distribution function for circular macromolecules with $n = 49$, which evolves as $\tilde{S}_1$ changes from 0.050 to 0.150 with step size 0.050.

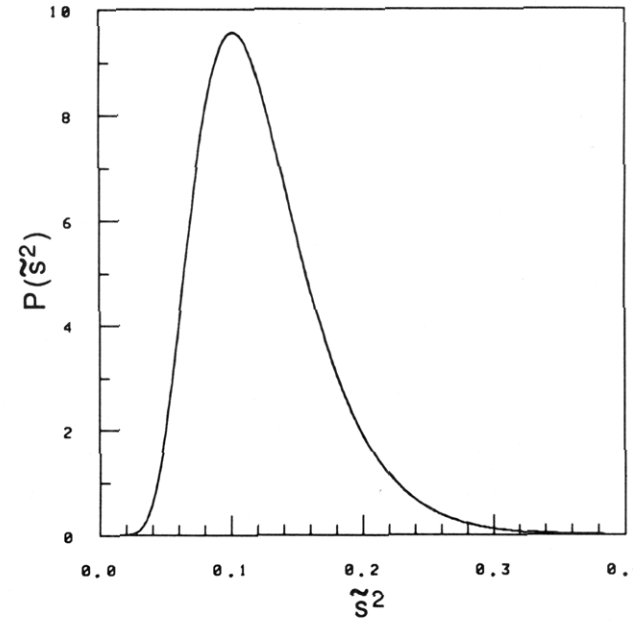

Figure 12. Distribution function of the squared radius of gyration as a contraction of the complete shape distribution function for circular macromolecules with $n = 11$. Also plotted is the same function as calculated from the existing analytic expression. The two curves are indistinguishable on this scale.

for 11-bead circular macromolecules. These quantities are of interest as elementary measures of the asymmetry of the distributions. It is noticed that all three distributions possess only one maximum, corresponding to the respective linear combination of the most probable values of the three principal components found in the shape distribution function. It is of particular interest to note that the most probable value of $\tilde{S}_1 - 2\tilde{S}_2 + \tilde{S}_3$ has a value slightly greater than zero, implying that $\tilde{S}_1 - \tilde{S}_2$ is apt to be slightly greater than $\tilde{S}_2 - \tilde{S}_3$, but these two differences are also fairly likely

Table II
Most Probable and Average Configurations (Asymmetric Ellipsoids)

n	$*\tilde{S}_1$	$*\tilde{S}_2$	$*\tilde{S}_3$	$P(*\tilde{S}_1,*\tilde{S}_2,*\tilde{S}_3)$	$\langle\tilde{S}_1\rangle$	$\langle\tilde{S}_2\rangle$	$\langle\tilde{S}_3\rangle$	R_1[a]	R_2[a]
				Circular					
11	0.050	0.02	0.01	44302.1	0.082	0.031	0.011	7.27	2.78
49	0.045	0.02	0.01	64369.3	0.079	0.030	0.011	7.12	2.70
				Linear					
11	0.08	0.03	0.01	8937.50	0.169	0.042	0.013	12.81	3.22
29	0.07	0.02	0.01	9883.62	0.167	0.042	0.013	12.41	3.11

[a] $R_j = \langle S_j\rangle/\langle S_3\rangle$.

Table III
Most Probable Configurations for Linear and Circular Chains for Selected Values of $\tilde{S}_1$

n	$\tilde{S}_1$	$*\tilde{S}_2$	$*\tilde{S}_3$	$P(\tilde{S}_1,*\tilde{S}_2,*\tilde{S}_3)$
		Circular		
49	0.050	0.020	0.010	63607.8
	0.100	0.020	0.010	23088.3
	0.150	0.020	0.010	5180.15
		Linear		
29	0.09	0.030	0.015	9395.11
	0.19	0.025	0.010	5343.76
	0.29	0.025	0.010	2377.24
	0.39	0.025	0.010	986.838

Table IV
Most Probable Configurations for Linear and Circular Chains for Selected Values of $\tilde{S}_2$

n	$\tilde{S}_2$	$*\tilde{S}_1$	$*\tilde{S}_3$	$P(*\tilde{S}_1,\tilde{S}_2,*\tilde{S}_3)$
		Circular		
49	0.020	0.045	0.010	64369.4
	0.040	0.065	0.010	17737.4
	0.060	0.085	0.010	3218.85
		Linear		
29	0.03	0.08	0.01	9628.37
	0.05	0.10	0.01	3906.28
	0.07	0.12	0.01	1365.51

Table V
Most Probable Configurations for Linear and Circular Chains for Selected Values of $\tilde{S}_3$

n	$\tilde{S}_3$	$*\tilde{S}_1$	$*\tilde{S}_2$	$P(*\tilde{S}_1,*\tilde{S}_2,\tilde{S}_3)$
		Circular		
11	0.040	0.080	0.050	91.2920
49	0.010	0.045	0.020	64369.3
	0.020	0.055	0.030	12802.0
	0.040	0.075	0.050	75.9735
61	0.040	0.080	0.050	76.2454
		Linear		
11	0.040	0.110	0.055	148.549
29	0.01	0.07	0.02	9883.62
	0.02	0.08	0.03	5053.78
	0.03	0.10	0.04	929.109
	0.04	0.11	0.05	158.197
	0.05	0.12	0.06	28.1085
	0.07	0.14	0.08	1.03998
61	0.04	0.11	0.05	159.135

Table VI
Most Probable and Average Dimensions for Selected Linear Combinations of $\tilde{S}_1$, $\tilde{S}_2$, and $\tilde{S}_3$ for 11-Bead Circular Chains

σ	$*\sigma$	$P(*\sigma)$	$\langle\sigma\rangle$	$\langle\sigma^2\rangle$
$\tilde{S}_1 + \tilde{S}_2 + \tilde{S}_3$	0.1000	9.55336	0.1240	0.0176
$\tilde{S}_1$	0.0600	12.6120	0.0821	0.0081
$\tilde{S}_2$	0.0225	35.7459	0.0303	0.0011
$\tilde{S}_3$	0.0075	73.3249	0.0105	0.0001
$2\tilde{S}_1 - \tilde{S}_2 - \tilde{S}_3$	0.0750	7.03132	0.1224	0.0198
$\tilde{S}_1 - 2\tilde{S}_2 + \tilde{S}_3$	0.0150	12.9103	0.0330	0.0026
$\tilde{S}_1 + \tilde{S}_2 - 2\tilde{S}_3$	0.0675	10.4519	0.0912	0.0103

to be found with equal values. Probability distributions for other quantities that depend nonlinearly upon three

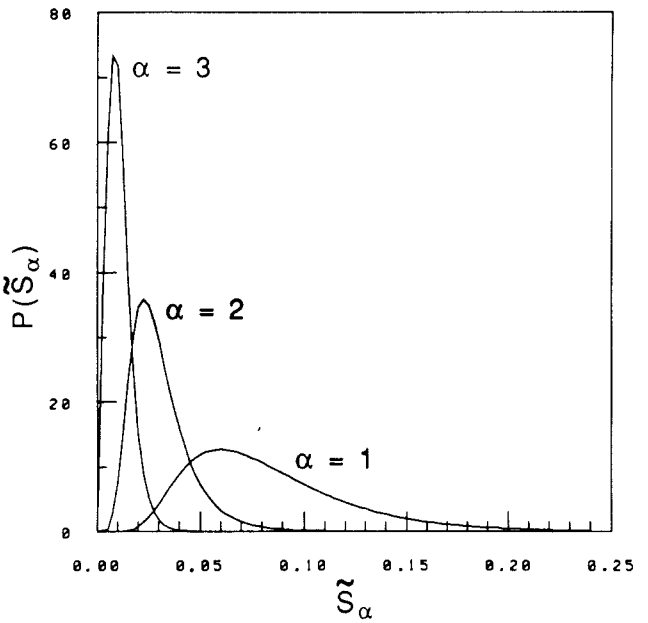

Figure 13. Distribution functions of individual principal components for 11-bead circular macromolecules.

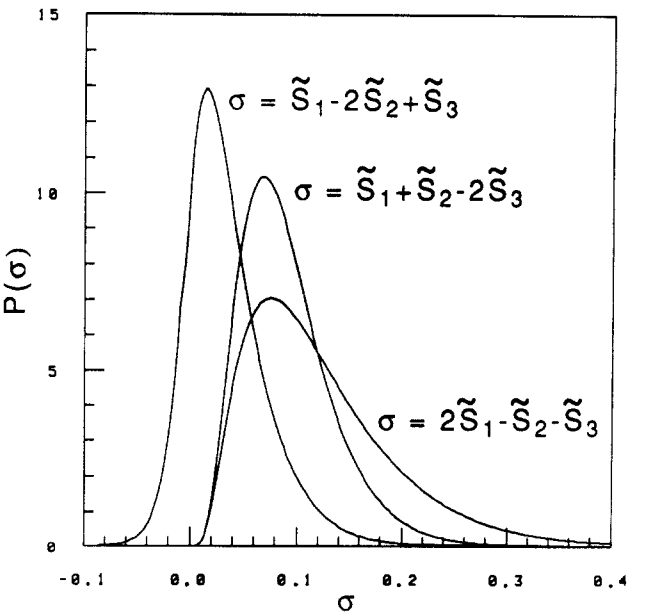

Figure 14. Distribution functions of three linear combinations of $\tilde{S}_1$, $\tilde{S}_2$, and $\tilde{S}_3$ for a circular macromolecule with 11 beads.

principal components, such as those that measure asphericity and prolateness,[8,9,31-33] may also be investigated by constructing integrals similar to that in eq 27.

In conclusion, three-dimensional shape distribution functions for both linear and circular macromolecules have been numerically evaluated. Several contracted distribution functions of linear combinations of the three principal components have also been evaluated for circular

macromolecules with $n = 11$. Simple asymptotic formulas for macromolecules that have force constant matrices in which the second smallest eigenvalue has odd degeneracy have been obtained. Molecules falling into this class are linear chains, star-shaped molecules, and random elastic networks. Two specific computational algorithms for linear and circular chains and a generic algorithm for molecules of arbitrary complexity have been developed, which make good use of the vector and parallel processing of IBM 3090 systems. These algorithms provide effective numerical methods for evaluating similar integrals that appear in the diverse mathematical disciplines of multivariate statistics, harmonic analysis, and number theory.[6,15,34,35] It is found that Gaussian macromolecules are highly anisometric when viewed from a coordinate frame that is aligned with the principal axes of inertia, that their most probable configurations are quite compact with respect to the mean-square dimensions, and that their extended configurations are essentially linear.

Acknowledgment. This work was supported by the Department of Energy through Grant DE-FG06-84ER45123. We thank both the IBM Corp. for a materials science grant leading to acquisition of an IBM 4381 and the University of Washington for donating computing time on its IBM 3090-200E. The part of this work that involves vector and parallel processing was conducted by using the Cornell National Supercomputer Facility (CNSF), a resource of the Cornell Theory Center, which is funded in part by the National Science Foundation, New York State, the IBM Corp., and members of the Center's Corporate Research Institute. The equal probability-density contour plots were made with use of the graphics package Insight of BIOSYM Technologies, Inc., by Rolland Thomas and B. E. Eichinger. G.W. thanks the CNSF for providing him an opportunity to attend the 1989 Cornell Advanced Summer Institute in Supercomputing.

References and Notes

(1) Šolc, K.; Stockmayer, W. H. *J. Chem. Phys.* **1971**, *54*, 2756.
(2) Šolc, K. *J. Chem. Phys.* **1971**, *55*, 335.
(3) Kuhn, W. *Kolloid-Z.* **1934**, *68*, 2.
(4) Eichinger, B. E. *Macromolecules* **1977**, *10*, 671.
(5) Eichinger, B. E. *Macromolecules* **1985**, *18*, 211.
(6) Muirhead, R. J. *Aspects of Multivariate Statistical Theory*; John Wiley & Sons: New York, 1982; Chapter 7.
(7) Theodorou, D. N.; Suter, U. W. *Macromolecules* **1985**, *18*, 1206.
(8) Rudnick, J.; Gaspari, G. *Science* **1987**, *237*, 384.
(9) Bishop, M.; Saltiel, C. J. *J. Chem. Phys.* **1988**, *88*, 3976.
(10) Fixman, M. *J. Chem. Phys.* **1962**, *36*, 306.
(11) Wei, G.; Eichinger, B. E. *Macromolecules* **1989**, *22*, 3429.
(12) Eichinger, B. E.; Shy, L. Y.; Wei, G. *Makromol. Chem., Macromol. Symp.* **1989**, *30*, 237.
(13) Wei, G. *J. Chem. Phys.* **1989**, *90*, 5873.
(14) Wei, G.; Eichinger, B. E. *J. Math. Phys.*, in press
(15) James, A. T. *Ann. Math. Stat.* **1964**, *35*, 475.
(16) Wei, G. *J. Math. Phys.*, in press.
(17) The IMSL Library. *Reference Manual*, 9th and 10th eds.; International Mathematical and Statistical Libraries: Houston, TX, 1982 and 1987; Programs MMDEI, MMBSIO, DQAWF, DBSIOE, and DQDAG.
(18) Gerald, C. F.; Wheatley, P. O. *Applied Numerical Analysis*, 3rd ed.; Addison Wesley: Reading, MA, 1984; Chapter 4.
(19) Wei, G. Ph.D. Thesis, University of Washington, 1990.
(20) Wei, G.; Eichinger, B. E. *J. Math. Phys.* **1990**, *31*, 1274.
(21) Coriell, S. R.; Jackson, J. L. *J. Math. Phys.* **1967**, *8*, 1276.
(22) Piessens, R.; deDoncker-Kapenga, E.; Uberhuber, C. W.; Kahaner, D. K. *QUADPACK—A Subroutine Package for Automatic Integration*; Springer-Verlag: New York, 1983.
(23) Gradshteyn, I. S.; Ryzhik, I. M. *Table of Integrals, Series, and Products*; Academic Press: New York, 1965; Entry 8.212.6, p 925; Entry 9.261.1, p 1067.
(24) Abramowitz, M.; Stegun, I. A. *Handbook of Mathematical Functions*; Applied Mathematics Section 55; National Bureau of Standards: Washington, DC, 1970; p 230.
(25) Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. *Numerical Recipes*; Cambridge University Press: Port Chester, NY, 1986; Chapter 4.
(26) Šolc, K. *J. Chem. Phys.* **1972**, *5*, 705.
(27) Šolc, K.; Stockmayer, W. H. International Symposium on Macromolecules, Helsinki, Finland, 1972; Preprint No. II-86.
(28) Shy, L. Y.; Eichinger, B. E. *Macromolecules* **1986**, *19*, 838.
(29) Zimm, B. H. *Phys. Rev. Lett.* **1988**, *61*, 2965.
(30) Eichinger, B. E. *Macromolecules*, in press.
(31) Aronovitz, J. A.; Nelson, D. R. *J. Phys. (Les Ulis, Fr.)* **1986**, *47*, 1445.
(32) Gaspari, G.; Rudnick, J.; Beldjenna, A. *J. Phys. A: Math. Gen.* **1987**, *20*, 3393.
(33) Aronovitz, J. A.; Stephen, M. J. *J. Phys. A: Math. Gen.* **1987**, *20*, 2539.
(34) Satake, I. *Notices Am. Math. Soc.* **1980**, *27*, 442.
(35) Hua, L. K. *Harmonic Analysis of Functions of Several Complex Variables in the Classical Domains*; Translation of American Mathematical Society: Providence, RI; 1963.